NASA Contractor Report 185670

NASA-CR-185670-VOL-1
19920012021

National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program—1991

Volume 1

William A. Hyman and Stanley H. Goldstein, Editors

Grant NGT 44-001-800

December 1991

NASA Contractor Report 185670

National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program—1991

Volume 1

William A. Hyman, Editor
Texas A&M University
College Station, Texas

Stanley H. Goldstein, Editor
University Programs Office
Lyndon B. Johnson Space Center
Houston, Texas

Grant NGT 44-001-800

National Aeronautics and Space Administration
Lyndon B. Johnson Space Center
Houston, Texas

December 1991

PREFACE

The 1991 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by Texas A&M University and JSC. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are

1. To further the professional knowledge of qualified engineering and science faculty members

2. To stimulate an exchange of ideas between participants and NASA

3. To enrich and refresh the research and teaching activities of participant's institutions

4. To contribute to the research objectives of the NASA Centers

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1991. Volume 1 contains reports 1 through 14, and Volume 2 contains reports 15 through 27.

This Page Intentionally Left Blank

CONTENTS

Volume 1

Volume 2

USE OF TAGUCHI DESIGN OF EXPERIMENTS TO DETERMINE ALPLS ASCENT DELTA-V SENSITIVITIES AND TOTAL MASS SENSITIVITIES TO RELEASE CONDITIONS AND VEHICLE PARAMETERS

Final Report

NASA/ASEE Summer Faculty Fellowship Program – 1991

Johnson Space Center

Prepared By:	Hector Ramon Carrasco
Academic Rank:	Assistant Professor
University and Dept.:	Florida International University Department of Industrial Engineering University Park Miami, Florida 33199
NASA/JSC	
Directorate:	Engineering and Development
Division:	Systems Engineering Division
Branch:	Systems Definition Branch
JSC Colleague:	Charles Mallini
Date Submitted:	August 13, 1991
Contract Number:	NGT 44-001-800

Use of Taguchi Design of Experiments to Determine ALPLS Ascent Delta-V Sensitivities and Total Mass Sensitivities to Release Conditions and Vehicle Parameters

ABSTRACT

NASA is currently assessing the engineering feasibility, safety and reliability, and the infrastructure and operational requirements of an of an air launched personnel launch system (ALPLS) for transportation of personnel to low earth orbit. One of the Study Requirements calls for the determination of ALPLS ascent delta-V sensitivities to release conditions (Altitude, Mach, Flight Path Angle, Delay Time) and vehicle parameters (Isp, Ignition Tvac/W, L/D).

Previous parametric studies of this type have been accomplished by varying and optimizing one parameter at a time. Parameter interactions are considered by re-evaluating sensitive parameters while holding other variables at or near previously determined optimal values. A traditional, scientific, approach results in the experimenter methodically evaluating every possible combination of the parameters. Note that if an experiment involved seven parameters, each with only two possible levels, a total of 2^7 or 128 experiments would have to be conducted. Statisticians, as advocated by Taguchi, have developed more efficient test plans, which are referred to as fractional factorial experiments (FFEs). Certain treatment conditions are chosen to maintain orthogonality among the various factors and interactions. Taguchi design of experiments methods can be used to more efficiently analyze the optimal values of the parameters indicated above. By using Fractional Experiment Designs (Orthogonal Arrays) the number of runs can be significantly reduced with a better evaluation of the parameter interactions. Orthogonal experiment design is a time-saving testing strategy that draws on orthogonal arrays to pinpoint areas where variation may be successfully reduced.

The objective of this study is to evaluate the use of Taguchi's Design of Experiment Methods to improve the effectiveness of this and future parametric studies. Taguchi Methods will be applied in addition to the typical approach to provide a mechanism for comparing the results and the cost or effort necessary to complete the studies. It is anticipated that results of this study should include an improved systematic analysis process, an increase in information obtained at a lower cost, and a more robust, cost effective vehicle design.

INTRODUCTION

The objective of this study is the demonstration of an improved methodology for the Parametric Design Phase of the space vehicle design process utilizing Taguchi Design of Experiment Methods. Although the design of experiment methods such as fractional orthogonal arrays are in common use in experimental settings, they are not being utilized in areas where computer models are available. Computer analysis of complex models can be more efficiently accomplished if computer runs are viewed and planned as experiments. Standardizing the evaluation process using design of experiment methods eliminates many of the common solution approaches such as random trial and error, the varying of one variable at a time, or the methodical and time consuming evaluation of every possible combination of variable levels. The use of experimental design should result in better designs using fewer computer runs while obtaining an interpretable evaluation of the designs sensitivity to their environment or more robust designs.

OVERVIEW OF PARAMETER DESIGN AND TAGUCHI METHODS

During product development, engineers and scientists are typically faced with two opposing requirements; improve or optimize performance and reduce or minimize cost. The process of searching for factors or parameters affecting performance and/or cost is usually experimental in nature. After a set of relevant parameters are assembled, the experimenter is faced with the task of determining which combination of parameter values achieve the desired results. It is the quality of this decision that can be improved when proper tests strategies are used. The most common test plan is to evaluate and optimize one parameter at a time. If there happens to be an interaction of the factor being studied with any other factor, the interaction will not be observed as all other factors are being held constant. This loss of interaction information is accepted as unavoidable as the only other perceived alternative is to perform a full factorial experiment, testing every possible combination of factors. This is usually not feasible as most real problems involve many parameters with three or more possible levels requiring thousands of experiment runs. Taguchi is an advocate of more efficient test plans, which are referred to as fractional factorial experiments (FFEs). FFEs use only a portion of the total possible combinations to estimate the main factor effects and some, not all, of the interactions (1) . FFEs are balanced experiments developed using orthogonal arrays. They can be used to evaluate many parameters with a minimum number of tests.

As suggested above, the goal of parameter design is to determine the parameter values of a product or process so that the product is functional, exhibits a high level of performance and is minimally sensitive to noise. The strategy is to design a high quality product which can be produced from low grade, low cost components with broad tolerances. This improved quality and reduced variability is achieved by selecting the optimum parameter values so that the product is least sensitive to input and noise variations. Conventional quality improvement techniques reduce product variation by removing the cause which is usually expensive, while Taguchi methods reduce variation by becoming less sensitive to input variations without removing the cause of variation. Improved quality is, therefore, achieved without or with minimal increase in cost.

Taguchi Methods take advantage of non-linear effects and the interaction between control factors and noise factors in order to obtain designs that are more "robust". Taguchi's approach to design of experiments utilizes techniques that are cost effective and directly applicable to the problems and requirements of modern industry. Although design of experiments has been around for some time, it is seldom used for product development.

A parameter design experiment typically involves two types of factors:

Control Factors whose levels can be set and maintained.

Noise Factors whose level either cannot or will not be set or maintained, yet which could affect the performance of the functional characteristics

Parameter design examines interactions between control factors and noise factors in order to achieve robustness. It is a search for parameter levels at which a characteristic is stable, despite the use of inexpensive components and materials or external conditions.

The major steps in designing, conducting, and analyzing an experiment are as follows:

1. Selection of factors and/or interactions to be evaluated
2. Selection of number of levels for the factors
3. Selection of appropriate Orthogonal Array
4. Assignment of factors and/or interactions to columns
5. Performance of experiments
6. Analysis of results
7. Performance of confirmation experiment.

The goal of the experiment is to determine the combination of parameters that gives the most stable and reliable performance at the lowest cost.

PROBLEM STATEMENT

The Systems Definition Branch, Systems Engineering Division, is currently assessing the engineering feasibility, safety and reliability, and the infrastructure and operational requirements of an Air Launched Personnel Launch System (Air Launched PLS) for transportation of personnel to low earth orbit. One of the study requirements calls for the determination of the ascent delta-V sensitivities to release conditions (Altitude, Mach, Flight Path Angle, Delay time) and vehicle parameters (Isp, Ignition Tvac/W, L/D).

Previous parametric studies of this type have been accomplished by varying and optimizing one parameter at a time. Parameter interactions are considered by re-evaluating sensitive parameters while holding other variables at or near previously determined optimal values. This approach, while time consuming, is preferred by analysts as it provides considerable insight regarding the sensitivities and costs associated with deviations from optimal conditions. This study will evaluate the use of Taguchi Design of Experiment Methods for space vehicle design parametric studies by performing an analysis of the Air Launched PLS in conjunction with the traditional approach to provide a mechanism for comparing the results and the effort necessary to complete the studies.

EXPERIMENTAL DESIGN AND ANALYSIS

Referring to the Air Launched Personnel Launch System study statement of work, section IIIB called for the following:

> "Determine ALPLS ascent delta-V sensitivities to release conditions (Altitude, Mach, Flight Path Angle, Delay Time) and vehicle parameters (Isp, Ignition Tvac/W, L/D). Sensitivities are to be evaluated over the following parameter ranges using a two-stage vehicle analysis:
>
> - Altitude: 25K to 45K ft.
> - Mach: 0.7 (or based upon aircraft performance)
> - Gamma: 0 to 30 deg.
> - Delay Time: 0 to 6 sec.
>
> - Isp: Based upon selected propulsion systems
> - Ignition Tvac/W: 1.3 to 1.9 (Ballistic
>
> 0.9 to 1.9 (Aero Assisted)
> - L/D: Ballistic, low, medium, high aero assist"

The process of designing, conducting, and analyzing the parametric study directly followed the steps given earlier.

Selection of Parameters and Levels to be Evaluated

The number of parameters given above was reduced to six with the conclusion that the aircraft would be flown at maximum velocity for the release of the PLS. The parameters and the levels selected for the Taguchi Study were as follows:

Variable	Levels		
Altitude:	25,000 ft.	35,000 ft.	45,000 ft.
Gamma:	0 degrees	10 degrees	20 degrees
Delay Time:	2 seconds	4 seconds	6 seconds
Isp:	LO_X/LH_2	$A50/N_2O_4$	HTBP
Ign Tvac/W	1.5	1.7	1.9
L/D	Ballistic	Winged	

The Flight Path Angle (gamma) was varied in 10 degree increments to conform with the increment size being used in the traditional approach. This limited the range to a total of 20 degrees in order to limit the number of levels in the Taguchi study to three. The Lift to Drag ratio (L/D) was reduced to two levels with project teams decision to model only one aero-assist version having wings sized to maintain a horizontal flight path and facilitate separation from the aircraft at the time of release.

Selection of Appropriate Orthogonal Array

The selection of an orthogonal array depends on the number of factors and interactions to be modeled and the number of levels for the factors. Initially, little was known about the relevant interactions; however, it was believed that the engine type (Isp) and L/D parameters were the dominant factors. The L18 orthogonal array, shown in Table I, was selected for the study as it can accommodate two and three level parameters and the interactions between the three level columns are distributed more or less uniformly to all of the other three-level columns (2).

Assignment of Factors to Columns

Table II contains the parameter column assignments. The Lift to Drag ratio was assigned to column one as it was the only two level parameter while the Isp was assigned to column two to enable the modeling of the interactions between the two probable dominant variables. It was believed that the prevalent interaction would occur between these two parameters due to the significant differences in the magnitude of the Isps. These column assignments also facilitated the setup for the various trials as different computer models were needed when switching from a ballistic vehicle to a winged vehicle and different rocket sizing models were needed for each different engine type. Columns seven and eight remained unassigned providing a measure of error. It should be pointed out at this time that the selection of the L18 orthogonal array was not appropriate for this study as the main effects of the first two columns and their interaction dominated the experiment. This can be readily seen in the raw data, also shown in Table II. The very large masses (m glow) of the Winged-HTBP vehicles significantly affected the results of the study to the extent that

Table I. L18 Orthogonal Array

$L_{18}(2^1 \times 3^7)$

Trial no.	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	3	1	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

Note: This is a specially designed array. An interaction is built between the first two columns. This interaction information can be obtained without sacrificing any other column. Interactions between three-level columns are distributed more or less uniformly to all the other three-level columns, which permits investigation of main effects. Thus, it is a highly recommended array for experiments.

Table II. Air Launched PLS Parameter Column Assignments

	L/D	Isp	Alt	Gamma	Delay	T/W	Error		Results	
	Column no.								Total ΔV	m glow
Trial no.	1	2	3	4	5	6	7	8	(ft/sec)	(lbm)
1	Ballistic	LOx/LH2	25	0	2	1.5	1	1	29,661	730,965
2	Ballistic	LOx/LH2	35	10	4	1.7	2	2	28,572	661,753
3	Ballistic	LOx/LH2	45	20	6	1.9	3	3	27,875	624,690
4	Ballistic	A50/N2O4	25	0	4	1.7	3	3	29,690	1,949,102
5	Ballistic	A50/N2O4	35	10	6	1.9	1	1	28,393	1,654,014
6	Ballistic	A50/N2O4	45	20	2	1.5	2	2	28,042	1,485,159
7	Ballistic	HTBP	25	10	2	1.9	2	3	28,417	2,349,550
8	Ballistic	HTBP	35	20	4	1.5	3	1	28,384	2,332,576
9	Ballistic	HTBP	45	0	6	1.7	1	2	28,369	2,324,922
10	Winged	LOx/LH2	25	20	6	1.7	2	1	27,929	722,221
11	Winged	LOx/LH2	35	0	2	1.9	3	2	27,808	773,352
12	Winged	LOx/LH2	45	10	4	1.5	1	3	28,309	904,770
13	Winged	A50/N2O4	25	10	6	1.5	3	2	28,157	1,914,820
14	Winged	A50/N2O4	35	20	2	1.7	1	3	27,777	2,028,850
15	Winged	A50/N2O4	45	0	4	1.9	2	1	27,781	2,467,980
16	Winged	HTBP	25	20	4	1.9	1	2	27,601	2,789,500
17	Winged	HTBP	35	0	6	1.5	2	3	27,839	3,502,700
18	Winged	HTBP	45	10	2	1.7	3	1	27,542	4,488,800

all of the parameter levels used in the 18th trial were found to be the worse possible settings in the data analysis phase although an altitude of 45,000 ft. and a delay time of two seconds were expected to be optimal.

Performance of the Experiment

The computer analysis necessary for the completion of the 18 trial runs was performed by the System Definition Branch of the Systems Engineering Division and their Engineering Support Contractor, Lockheed Engineering and Science Company. The parametric study using the traditional approach required approximately 90 computer program setups, each requiring approximately 30 computer runs for a total number of computer runs in the order of 2500 to 3000. The Taguchi Method required only 18 program setups with only a few computer runs to determine optimal delta-V splits and flight profile to minimize total vehicle mass at the specified initial parameter conditions. At this point the effort required to perform the Taguchi study is significantly less than the traditional approach; however, more analysis will be necessary before the Taguchi Method provides the information required for the Air Launched PLS study.

Analysis of Results

Analysis of variance is typically used to determine the extent to which factors contribute to variation and to test significance of factorial effects. The factors that strongly effect variation can, however, be determined by recognizing the differences in the average effects of factors. These differences in effect are clearly seen in response graphs (3).

The Response Graphs, shown in Figure 1, show the response of the gross lift off mass with respect to each of the six parameters studied. The graphs show a slight, possibly significant preference to the ballistic configuration and a strong preference for the LO_X/OH_2 fueled engine. The remaining parameters appear to have little effect on the vehicle mass, with the optimal release conditions at 25,000 ft altitude, a flight path angle of 20 degrees, a drop time of six seconds prior to firing of the engines and an optimal thrust to weight ratio of 1.9. The preference for a low release altitude can be explained by the realization that the winged space craft requires a significantly larger wing to support the craft in the thinner atmosphere of the higher altitude. The preference for large drop times is caused by the insensitivity of the winged vehicles to drop time as the available lift minimized the loss in altitude over time and the dominance of the winged vehicles due to their larger masses.

The Interaction Graphs, given in Figure 2, demonstrate the interactions between the Lift to Drag parameter and the other five parameters. Although the L18 orthogonal array is only designed to model the interaction between the first two columns, L/D versus Isp, considerable insight can be obtained from the study of all of the graphs. The L/D versus Isp graph shows that an interaction is present as the lines are not parallel. It should also be noted that both ballistic and winged vehicles preferred the LO_X/OH_2 engines. The L/D versus Altitude graph shows the preference for a high altitude release for the ballistic vehicle and a low altitude release for the winged space vehicle. The remaining graphs are more severely hampered by the interactions not modeled with unexplained switches in the preferences of the ballistic and winged vehicles. The Interaction Graphs, given in Figure 3, show a much more erratic behavior. The Altitude versus L/D graph again clearly shows the preference of a low altitude release for the winged vehicle and a high altitude release for the ballistic one. However, the Altitude versus Flight Path Angle, Altitude versus

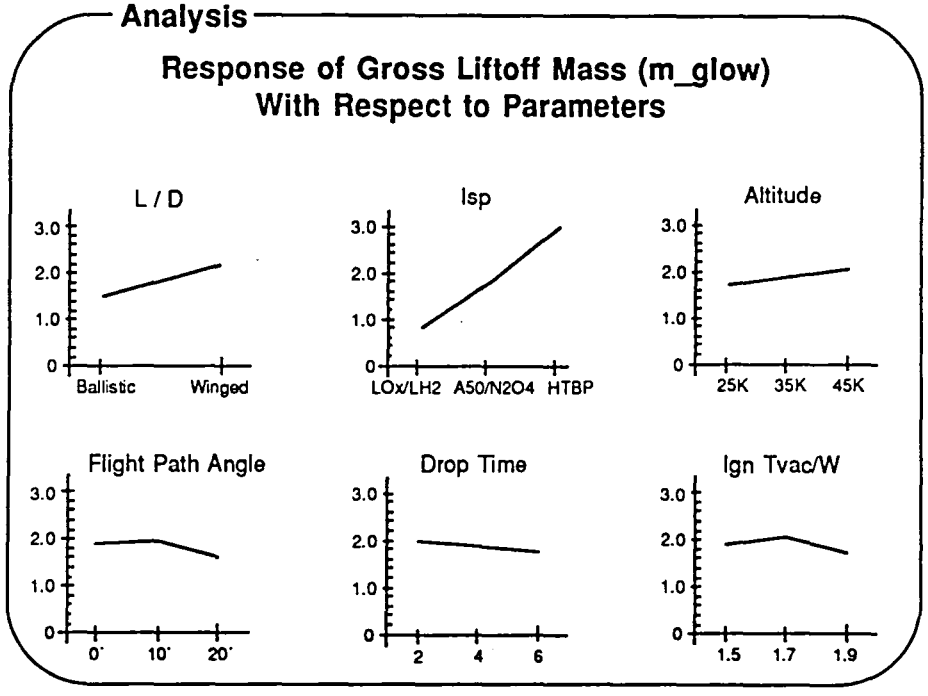


Figure 1. Air Launched PLS Response Graphs

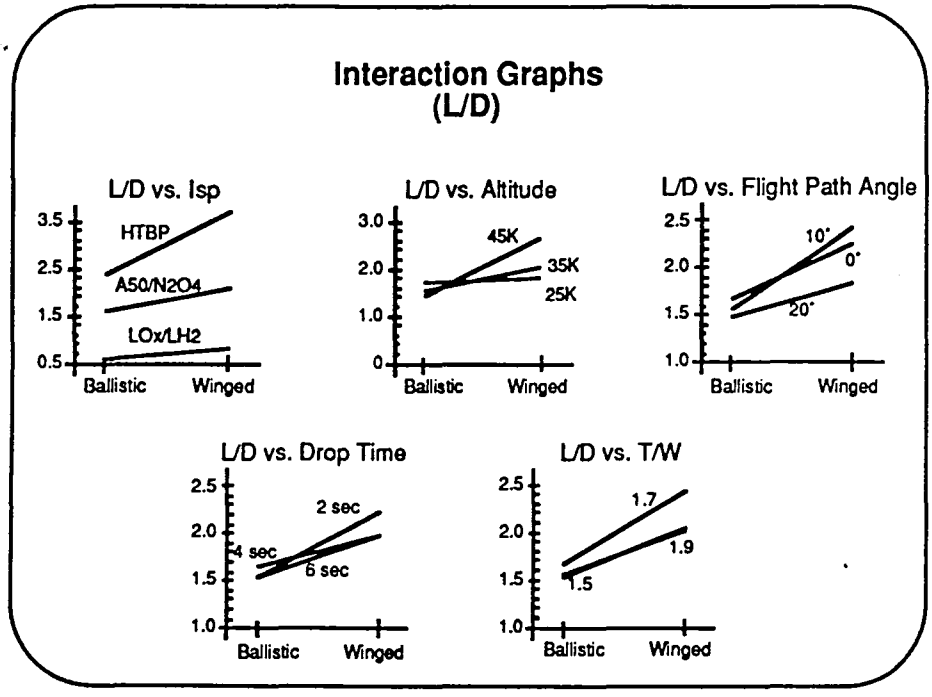


Figure 2. Lift to Drag Interaction Graphs

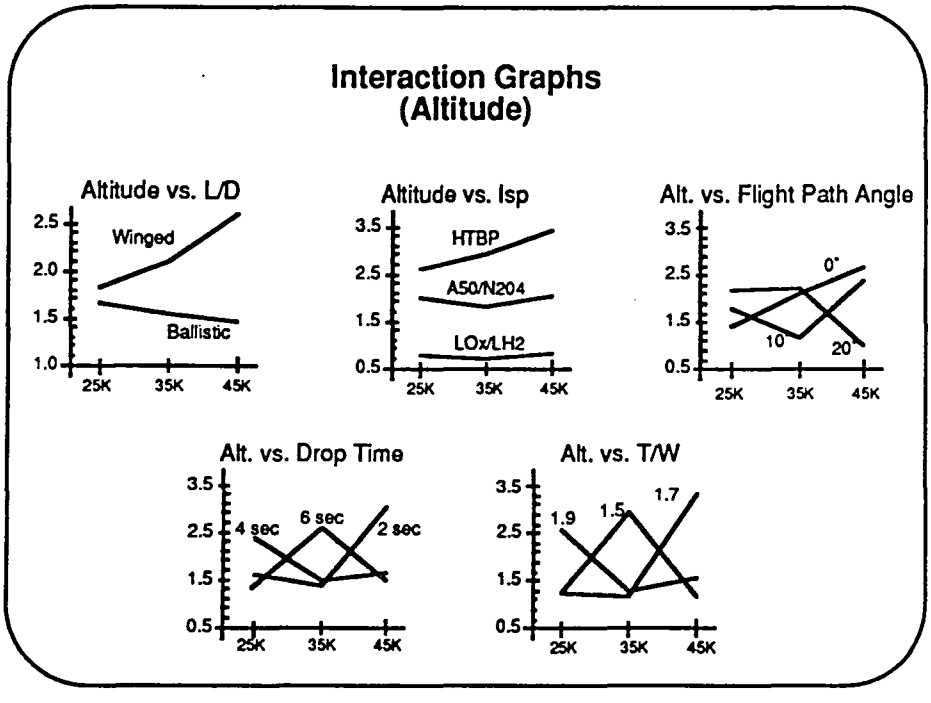


Figure 3. Altitude Interaction Graphs

Drop Time, and Altitude versus Thrust to Weight Ratio graphs only demonstrate the significance of the interactions. The primary conclusion that can be drawn from these graphs is that the L18 orthogonal array was not suitable for this study as the interactions were not as "uniformly distributed" as originally anticipated.

Another indication regarding the significance of the interactions can be seen in the analysis of variance, refer to Table III. The sum of squares for the factors in columns one and two are clearly more significant than the remaining factors. It is estimated that L/D, Isp, and their interaction account for as much as 90% of the total observed variation. It should be noted, however, that an analysis of variance is not valid, as an estimate of experiment error cannot be determined unless the experiment is replicated (4). As this experiment consist of computer runs, a repetition of the entire set of conditions would result in exactly the same values containing the same errors due to the simplifying assumptions and convergence tolerances.

RECOMMENDATIONS

The L/D - Isp interaction problem can be circumvented by studying the ballistic and winged vehicles separately. An L27 orthogonal array has been selected for further study as the anticipated interactions between Isp and the remaining parameters can be evaluated as shown in Table IV. The same 27 trials will be repeated for both L/D configurations. This will require an additional 54 computer setups and enough computer runs to minimize vehicle mass at the specified parameter conditions.

Table III. ANOVA for Air Launch PLS

	L/D	Isp	Alt	Gamma	Delay	T/W	Error		Results	
	Column no.									
Trial no.	1	2	3	4	5	6	7	8	Total ΔV	m glow
1	Ballistic	LOx/LH2	25	0	2	1.5	1	1	29,661	730,965
2	Ballistic	LOx/LH2	3	10	4	1.7	2	2	28,572	661,753
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	Winged	HTBP	35	0	6	1.5	2	3	27,839	3,502,700
18	Winged	HTBP	45	10	2	1.7	3	1	27,542	4,488,800

Level 1	14.113	4.418	10.456	11.749	11.857	10.870	L/D vs. Isp interactions	
Level 2	19.592	11.499	10.953	11.973	11.106	12.176	Ball-LOx/LH2	2.017
Level 3		17.788	12.296	9.983	10.742	10.659	Ball-A50/N2O4	5.088
Totals	33.705	33.705	33.705	33.705	33.705	33.705	Ball-HTBP	7.007
							Wing-LOx/LH2	2.400
SS	1.668	14.915	.302	.396	.108	.225	Wing-A50/N2/O4	6.411
v	1	2	2	2	2	2	Wing-HTBP	10.781
V	1.668	7.458	.151	.198	.054	.112	SS	17.604
							v	2
F	4.57	20.42	.41	.54	.15	.31	V	8.802
							F	24.10

Table IV. L27 Orthogonal Array

ORTHOGONAL ARRAY RECOMMENDED FOR FURTHER STUDY

	Isp	T/W		Alt		Gamma			Delay			L/D	
	Column no.												
Trial no.	1	2	3	4	5	6	7	8	9	10	11	12	13	Ballistic	Winged
1	LOx/LH2	1.5	1	1	25K	1	1	0°	1	1	2	1	1		
2	LOx/LH2	1.5	1	1	35K	2	2	10°	2	2	4	1	1		
⋮	⋮	⋮			⋮			⋮			⋮				
26	HTBP	1.9	2	1	35K	1	3	0°	3	2	6	2	1		
27	HTBP	1.9	2	1	45K	2	1	20°	1	3	2	3	2		

An L27 (3^{13}) array is recommended for further study with the Ballistic and Winged studies to be done separately. The L27 array was selected as it allows for the modeling of the interactions between Isp and each of the other parameters. The interactions between Isp and T/W will be modeled by columns 3 & 4, interactions between Isp and Altitude in columns 6 & 7, etc. This design should accurately evaluate the main effects (T/W, Altitude, Flight Path Angle, and Drop Time) as well as their interactions with Isp at a cost of requiring a total of 54 trials.

If this new experiment design successfully obtains the information required by the study's original statement of work, a total of 72 computer setup will have been required. Although the Taguchi Method is approaching the number of setups used in the traditional approach (90), a smaller number computer runs will be required using significantly less computer time than the traditional parametric study.

CONCLUSIONS

The application of Taguchi's Design of Experiment Methods to improve the effectiveness of complex computer models is a new learning process subject to a learning curve. A common initial strategy requires the evaluation of a large number of parameters using a saturated array. This assists the analyst in determining which factors are relevant for further experiments. Knowledge of the system being modeled greatly reduces the initial effort required by improving the process of determining the proper parameters. The initial selection of the L18 matrix was made with the knowledge that interactions existed, but without anticipating their significance. The project is currently in the second phase with an experimental design better suited to the interactions found.

As the study is not complete the anticipated results have not been achieved. The interim conclusions are as follows:

Anticipated results:

* Improved systematic analysis process -- While Taguchi's design of experiment methods are clearly more systematic, this project has not demonstrated the methods effectiveness.

* More information obtained at lower costs -- At this point in the project Taguchi Methods have not obtained the information required for the Air Launched PLS study.

* More robust, cost effective designs -- Sensitivity to noise and small variations in control parameters have not been addressed at this time.

Actual results:

* Considerable experience has been gained in the application of Taguchi Techniques in non-manufacturing environments.

* The first point in the learning curve has been obtained.

REFERENCES

1. Ross, P.J.: Taguchi Techniques for Quality Engineering. McGraw-Hill, 1988.

2. Taguchi, G.; Konishi, S.: Orthogonal Arrays and Linear Graphs: Tools for Quality Engineering. American Supplier Institute, Inc. (Dearborne), 1987.

3. Center for Taguchi Methods and Quality Function Deployment: Taguchi Methods Introduction to Quality Engineering, Workshop Course Manual. American Supplier Institute, Inc. (Dearborne), 1989.

4. Miller, I.; Freund, J.E.: Probability and Statistics for Engineers. Prentice-Hall, 1985.

OPTICAL CORRELATION

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1991

Johnson Space Center

Prepared By:	Steven S. Cotariu
Academic Rank:	Assistant Professor
University & Department:	United States Military Academy Department of Physics West Point, New York 10996
NASA/JSC	
Directorate:	Engineering
Division:	Tracking and Communications
Branch:	Tracking Techniques
JSC Colleague:	Richard D. Juday, Ph.D.
Date Submitted:	August 9, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

Pattern recognition may supplement or replace certain navigational aids on spacecraft in docking or landing activities. The need to correctly identify terrain features remains critical in preparation of autonomous planetary landing. One technique that may solve this problem is optical correlation. Correlation has been successfully demonstrated under ideal conditions; however, noise significantly affects the ability of the correlator to accurately identify input signals. Optical correlation in the presence of noise must be successfully demonstrated before this technology can be incorporated into system design.

An optical correlator is designed and constructed using a modified 2f configuration. Liquid crystal televisions (LCTV) are used as the spatial light modulators (SLM) for both the input and filter devices. The filter LCTV is characterized and an operating curve is developed. Determination of this operating curve is critical for reduction of input noise. Correlation of live input with a programmable filter is demonstrated.

INTRODUCTION

The ability to safely perform an autonomous planetary landing is critical in NASA's continuing exploration of our solar system. Round trip communication transmission from Earth to Mars require tens of minutes, depending upon the relative positions of the planets in their orbits. An unmanned craft must be independently capable of terrain identification and hazard avoidance. Pattern recognition techniques offer a possible solution to this problem.

Several approaches to this problem are currently under investigation at NASA. These include fuzzy logic, neural networks, fractal characterization and optical correlation. Although these approaches are at various stages of development, the optimum approach, or combination of approaches, has yet to be determined.

Optical correlation offers two distinct advantages over other more computationally intensive techniques: speed and weight. The essence of the optical correlator, the lens, computes the Fourier transform of an image in real time. Proper filtration of that transform and re-transformation with another lens allows for near real time recognition of visual scenes.

BACKGROUND

The origins of Fourier optics can be traced to Abbe and Porter in 1893 and 1906, respectively. Their efforts involved the modification of image information through spatial filtering in the Fourier plane. In 1931 Goldberg was granted a U.S. patent for his application of Fourier optics to character recognition.[1]

A textbook example of an optical correlator is the "4 f correlator." This configuration consists of two lenses, an input plane, filter plane and correlation plane. For lenses of focal length f, the total length of the correlator is four times f (figure 1). For a given object in the input plane and the corresponding filter in the filter plane a correlation peak will be detected in the correlation plane.

Early work with the 4 f correlator used photographic plates as input images. Matched filters were developed for each image and these filters were also placed on photographic plates. Although correlators of this type can adequately recognize a letter, they are sensitive to size and orientation variations.

Detection limitations due to this sensitivity have led to increased investigation in filter design. Variations of the matched filter include phase, amplitude and binary filters.

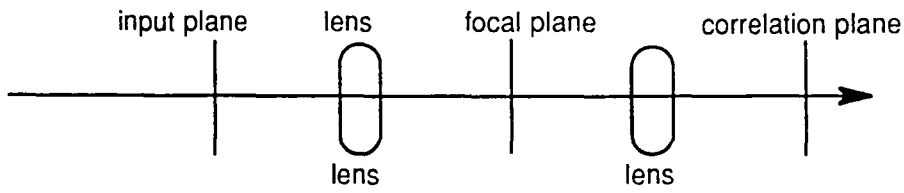


Figure 1. Collimated light enters from the left. The lenses are two focal lengths apart. The input, focal and correlation planes are one focal length from the nearest lens.

Ultimately the filter design must accommodate any limitations of the filter medium, the spatial light modulator (SLM). An SLM modulates amplitude, phase, or a combination of both amplitude and phase. The earliest SLMs were light blocks and photographic film. Deformable mirror devices (DMDs) have been used in the modulation of light. Low cost and ready availability have recently increased interest in the liquid crystal televisions (LCTVs) as a suitable SLM in either phase or amplitude mode.

Though any physical device is constrained to an operating curve and never truly has access to every point of a region on the complex plane (fully complex modulation), filter design and filter characterization have focused on either phase or amplitude and not both. Juday has developed an algorithm to select the point on a device operating curve that best incorporates the phase and amplitude information of the desired point on the complex plane.[2]

Other authors have noted both phase and amplitude modulation in LCTVs; however, none have attempted to operate in both phase and amplitude modulation.[3] With Juday's work a significant increase in signal to noise ratio (SNR) may be obtained by operating in phase and amplitude modulation.

In addition to developing a programmable filter on an LCTV, our research goals included providing live input to the correlator. A second LCTV, driven by an external camera, provided a logical course of action. In this configuration input images could easily be varied and filters could be rapidly changed.

EXPERIMENTAL CONFIGURATION

The LCTV uses a twisted nematic that rotates the polarization of incident light. The amount of rotation is dependent upon the alignment of the liquid crystal molecules. Molecular alignment is controlled through application of a voltage across the LCTV. The brightness control knob on the

Epson projector applies a bias voltage between the two surfaces of the LCTV. The potential difference across individual pixels is controlled via gray scale (0-255) application of additional voltages.

In normal operation, a dichroic sheet polarizer is placed in front of the LCTV and another crossed polarizer is place in back of the LCTV. The amount of polarization rotation induced by the LCTV then determines the total throughput.

The LCTVs used were removed from an Epson Crystal Image Video Projector. This projector uses three LCTVs; red, green and blue. The red LCTV was selected to serve as the filter and the green was selected to serve as the input SLM. Each of these LCTVs was mounted onto conventional optical mounts. Three foot extension cables were fabricated for ease of mounting. For convenience, the input and filter LCTVs were placed orthogonally in the optical path.

The service manual accompanying the projector reported the LCTVs to consist of 70400 pixels (320x220) arrayed in a 1.27 inch diagonal area with an aspect ratio of 4(H):3.1(V). The pixels were in rows and columns uniformly spaced 0.008 cm. and 0.009 cm. respectively. Filters were to be sent to the Epson via a PIP-1024-A Frame Grabber/Buffer. Digital images, 512 x 480 on the PIP, were converted to analog video, sent to the Epson and then converted back to digital and displayed on the LCTV. Careful analysis of the data transfer from the PIP to the LCTV revealed that the 480 rows were sent to the 220 LCTV rows in the following manner: the top 19 rows were not sent to the LCTV, the next 440 rows were sent in an overlapping manner; that is, the first PIP row was written on the first LCTV row, then the second a PIP was written on top of the first LCTV row. This was repeated over the entire LCTV surface, and the bottom 21 rows were not sent to the LCTV. From left to right, the first 13 rows on the PIP were not sent to the LCTV, the next 493 were sent in a fractional manner with 1.5085 PIP columns transferring to 1 LCTV column; the final six columns on the PIP were not sent to the LCTV.

In normal operation of the LCTV 0 gray scale corresponds to black (minimum transmission) and 255 gray scale corresponds to white (maximum transmission). The Epson service manual states that the minimum contrast ratio for each pixel will be 40:1. We found a contrast ratio of 63:1 on the LCTV.

We chose to operate our input LCTV in the conventional mode; however, we decided to run our filter in a white-black-white configuration. The most satisfactory results were found with the corresponding gray scales of 0-127-255. A maximum contrast ratio of 100:1 was found. The data are plotted in figure 2.

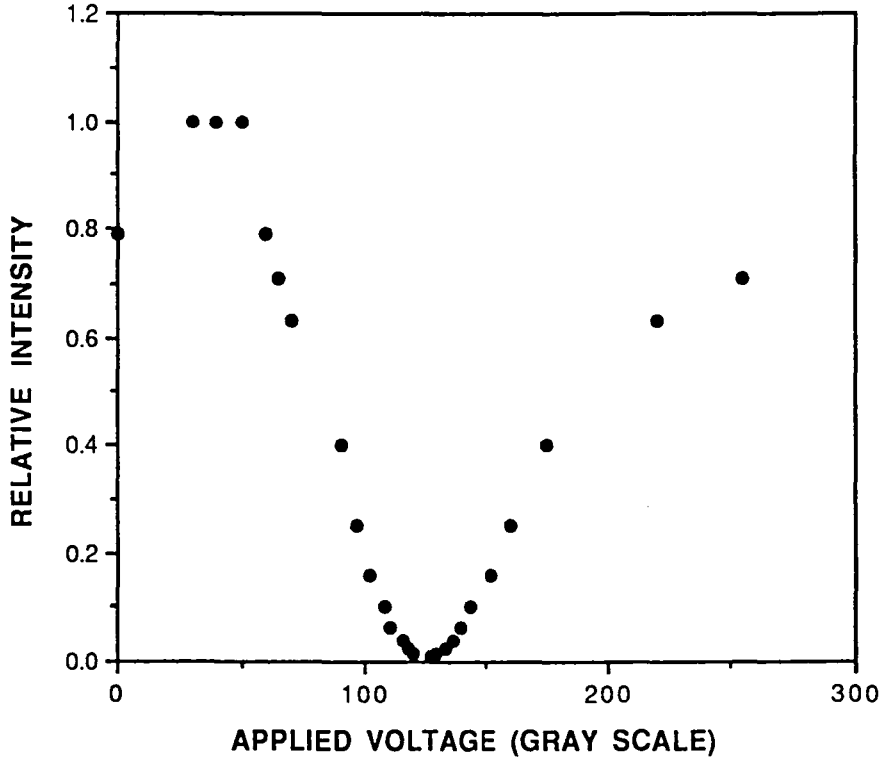


Figure 2. Intensity as a function of voltage

A procedure was developed for determining the optimum setting of the three control variables: input polarizer, output polarizer and brightness knob. The LCTV was divided into three sections, each with a different gray scale setting. The edge sections were set to 0 and 255, respectively, and the center section was set to 127. A trace of the intensity, running across each of the three sections, was displayed on a scope. The input polarizer was rotated until the intensity of the 0 and 255 was even. Using the output polarizer the center section intensity was minimized. Further minimization of the center section was accomplished by adjustment of the brightness knob. These three adjustments were then systematically repeated, minimizing the intensity of the center section, while maintaining equal intensity on the end sections. After the settings had been determined and recorded, all of the adjustment knobs were randomly set and the procedure was repeated several times to insure that the optimum set of adjustments had been determined.

Phase variation as a function of gray scale was determined using a Mach-Zehnder interferometer. Measurements were made with the LCTV and both polarizers in place, at the settings determined during the contrast optimization procedure. To make the phase measurements the LCTV was divided into two sections on one side the gray scale was set at 255 and the other side was interactively varied in increments of 5. The fringe variation was observed the entire time, noting a smooth transition, and data points were recorded every 1/4 fringe. Although only five data points were recorded, the curves displayed are accurate descriptions of the phase variation as a function of voltage (figure 3).

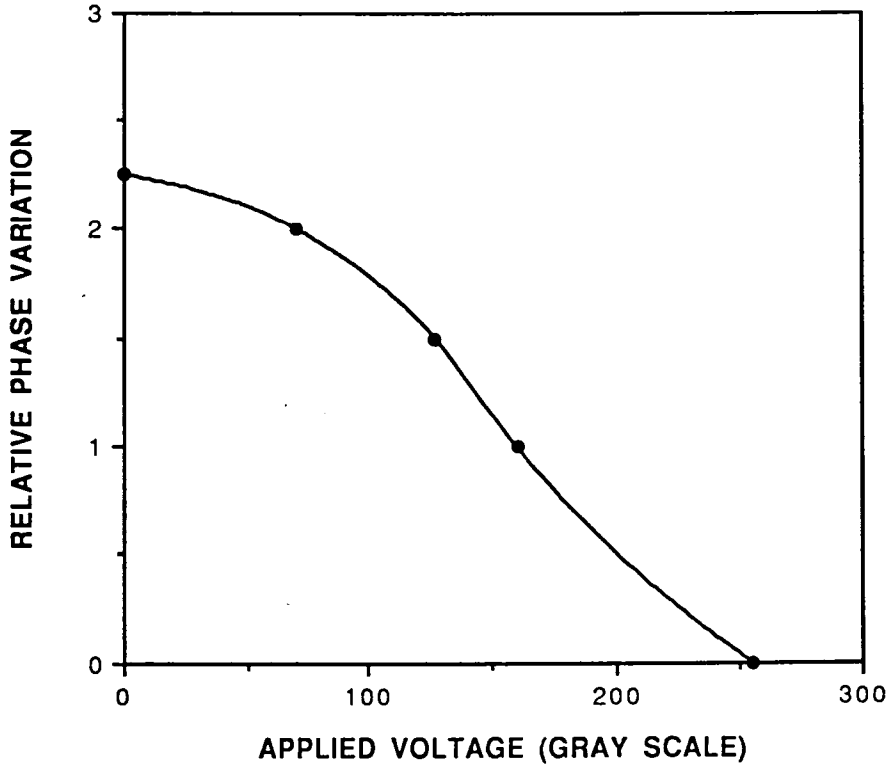


Figure 3. Phase as a function of voltage.

The coupled curve was then modeled using the five phase data points and their corresponding intensity measurements. Amplitude (E) was modeled as though the polarization rotation

leading to its variation were affine with phase, with

$$E=A\cos(k\theta-\beta) \qquad (1)$$

for A=0.9803, k=0.4578 and β=0.5295.

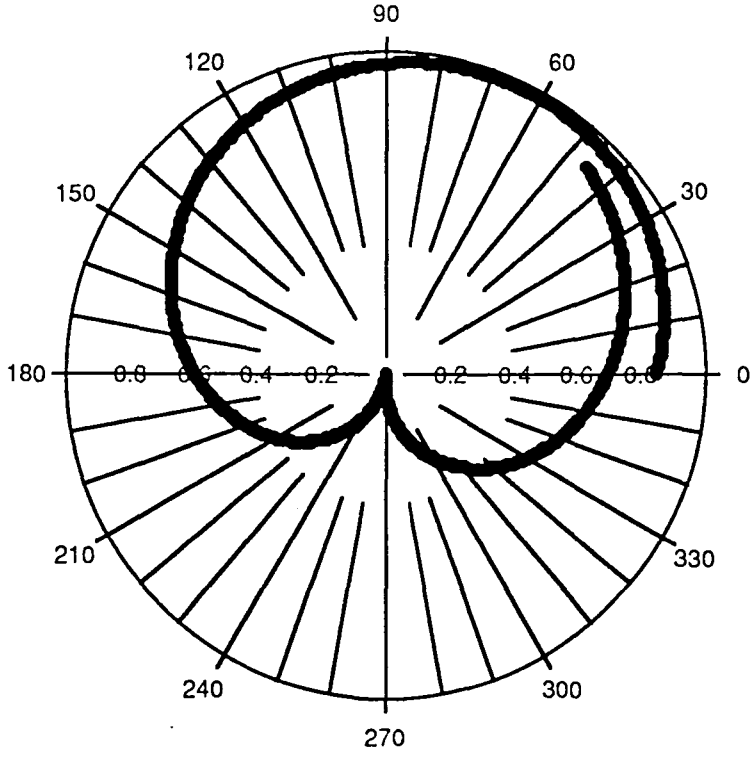


Figure 4. Operating Curve for maximum contrast.

The operating curve of the LCTV (phase vs intensity) is at figure 3. It is this curve that is used in the selection of optimal filter values. Corresponding voltage for any point on the operating curve can readily be found from Table 1.

TABLE 1. PHASE, AMPLITUDE AND VOLTAGE RELATIONSHIPS

$\theta=2.2652-3.7813 \times 10^{-3}V-2.0530 \times 10^{-5}V^2$

$I=-1.7090+1.6889 \times 10^{-2}V-2.8835 \times 10^{-2}V^2$ for $V>127$

$I=0.78755+1.6658 \times 10^{-2}V-2.9384 \times 10^{-4}V^2-9.9441 \times 10^{-8}V^3 +7.8447 \times 10^{-9} V^4$ otherwise

$A=0.89283+7.1321 \times 10^{-3}V-1.1596 \times 10^{-4}V^2$ otherwise

$A=-4.0612+5.3474 \times 10^{-2}V-1.9376 \times 10^{-4}V^2+2.3299 \times 10^{-7}V^3$ for $V>127$

$V=102.29+288.48A-687.5A^2+656.58A^3$ for $V>127$

$V=116.82-11.796A-56.913A^2$ for $40<V\leq127$

$V=125.91+245.35I-605.85I^2+725.02I^3$ for $V>127$

$V=116.96-70.353I-5.9588I^2$ otherwise

$V=251.08-46.990\theta-27.495\theta^2$

Where A is the amplitude (relative, scaled from 0 to 1, ±0.05), I is the square of the amplitude, V is the gray scale voltage (±2.5) and θ is the relative phase ($\pm0.12\ \pi$).

Of additional importance in characterization of LCTVs is the orientation of the director. The director is the longitudinal axis of the molecular dipole at the front surface of the LCTV. To determine the orientation of the director linearly polarized light is sent into the LCTV (with no voltage applied) and the polarization of the throughput is analyzed. The input polarization is systematically rotated and the throughput analyzed until settings are determined where the input and throughput light are both linearly polarized. This will only occur when linearly polarized light that is either perpendicular or parallel to the director enters the LCTV. At any other orientation, linearly polarized light will be converted to elliptically polarized light after passing through the LCTV.

Linearly polarized light that enters the LCTV perpendicular to the director will not experience a phase shift, regardless of voltage applied to the LCTV, whereas, linearly polarized light that enters the LCTV parallel to the director will experience a phase shift as varying voltages are applied to the LCTV. This information, combined with the settings determined above, can be used to uniquely determine

the orientation of the director. For the filter LCTV the director was found to be 73 degrees from the vertical on the LCTV.

Optimum lens selection for a 4 f correlator requires frequency matching of the input and filter SLMs. For a pixelated input SLM, optimum filtering occurs when one diffraction order from the input image just fits onto the filter. This occurs when:

$$f=Y/d\lambda \qquad (2)$$

where f is the transform lens focal length, Y is the diameter of the filter, d is the Nyquist frequency of the input SLM and λ is the wavelength of light used.[1]

For a 128 x 128 input image, a 128 x 128 phase only filter was calculated and sent from the PIP to the filter LCTV. Using the PIP to LCTV mapping determined earlier and equation 1, an appropriate focal length of 195 cm. was selected.

A 195 cm. lens was not available; had it been a 4f correlator it would then extend nearly eight meters, an unacceptable long length for our laboratory. To overcome these two problems a 2f modified correlator was designed using the lenses available in the lab (figure 5).[4]

A step by step procedure was developed for lens selection and placement in development of the 2f modified correlator. This procedure is presented below.

1. Lens 1 is selected by image pixel spacing, width of filter and wavelength (eqn 2.).

2. Input plane and lens 2 are selected for convenience and image size.

3. Lens 3 is selected such that $1/f_3=1/d+1/f_2$, where d is the distance from the input plane to the filter plane.

4. The correlation plane is positioned a distance s' from the filter plane such that $1/f_3=1/d+1/s'$.

5. Lens 4 is selected such that $1/f1=1/f4+1/x$, where x is the distance from the spatial filter to lens 4.

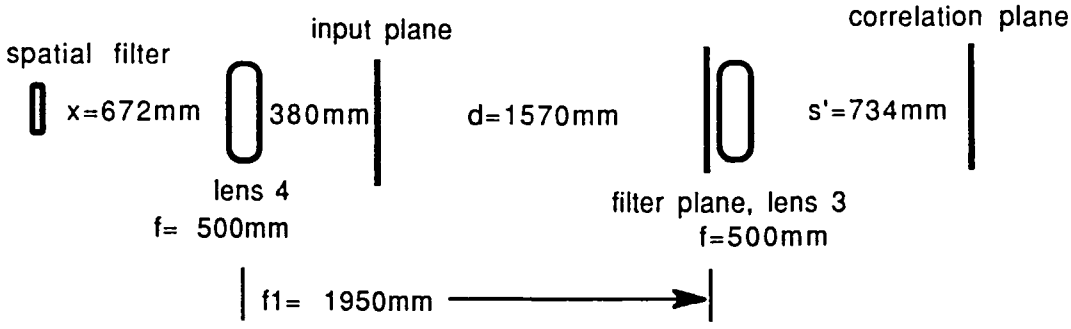


Figure 5. Final correlator configuration.

CONCLUSIONS

Procedures for characterization of LCTVs necessary to determine phase and amplitude cross coupling were developed. These procedures can be applied for optimization of filter design. An Epson-Epson correlator, with video input and phase only filter operated successfully on simple geometric objects. Correlation of live input was accomplished with a phase only filter displayed on the filter SLM. SNR in excess of 25:1 were recorded (figure 6).

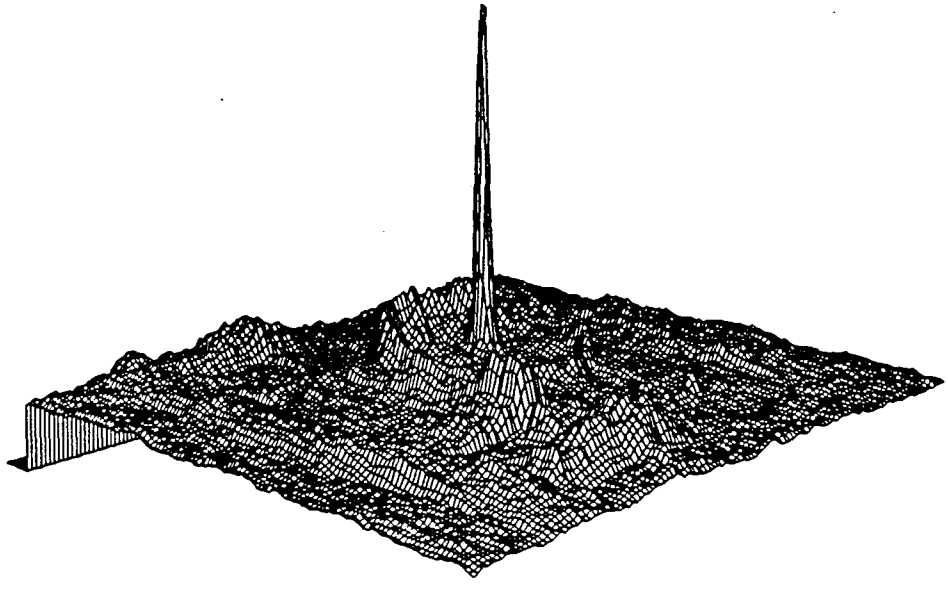


Figure 6. Intensity distribution in the correlation plane.

REFERENCES

1. J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, 1968).

2. R. Juday, "Correlation with a spatial light modulator having phase and amplitude cross coupling," Applied Optics **28**, 4865-4869 (1989).

3. D. Gregory et al., "Using the hybrid modulating properties of liquid crystal television," Applied Optics **30**, 1374-1378 (1991).

4. K. Fielding et al., "Optical correlator architectures re-examined," SPIE **1295**, 59-66 (1990).

Investigation of Techniques for Simulating Communications and Tracking Subsystems on Space Station Freedom

Final Report

NASA/ASEE Summer Faculty Fellowship Program -- 1991

Johnson Space Center

Prepared by: Louis A. DeAcetis, Ph.D.

Academic Rank: Professor of Physics

University & Department: Bronx Community College/CUNY
Department of Physics and Astronomy
Bronx, NY 10453

NASA/JSC

Directorate: Engineering

Division: Tracking and Communications

Branch: Systems Engineering

JSC Colleague: Oron Schmidt

Date Submitted: August 16, 1991

Contract: NGT-44-001-800

Abstract

The need to reduce the costs of Space Station Freedom has resulted in a major redesign and downsizing of the Station in general, and its Communications and Tracking (C&T) components in particular. Earlier models and simulations of the C&T Space-to-Ground Subsystem (SGS) in particular are no longer valid. There thus exists a general need for updated, high fidelity simulations of C&T subsystems. This project explored simulation techniques and methods that might be used in developing new simulations of C&T subsystems, including the SGS. Three requirements were placed on the simulations to be developed:

a) They run on IBM PC/XT/AT compatible computers

b) They be written in Ada as much as possible

c) Since control and monitoring of the C&T subsystems will involve communication via a MIL-STD-1553B serial bus, that the possibility of commanding the simulator and monitoring its sensors via that bus be included in the design of the simulator.

The result of the project is a prototype of a simulation of the Assembly/Contingency Transponder of the SGS, written in Ada, which can be controlled from another PC via a MIL-STD-1553B bus.

Introduction

The design of Space Station Freedom (SSF) has recently undergone a major redesign and reconfiguration. This has significantly impacted upon the Communications and Tracking Subsystem (CTS) and its space-to-ground (SGS) components. An earlier design was simulated [1] and used to test various aspects of the inter-system communications design of earlier models of SSF (so-called Demonstration 3B, [2]). These simulations are currently outdated, and the purpose of this investigation was to determine appropriate programming techniques that could be used in developing an updated simulation of the CTS in general, and the space-to-ground system in particular.

The various subsystems on SSF will be connected by two networks, one a token ring fiber optic network employing the Fiber Distributed Data Interface (FDDI) protocol, and other local "networks" consisting of MIL-STD-1553B busses to be used to control and monitor the operation of the component elements (generally comprised of what are called orbital replacement units, or ORU's) of the various subsystems on SSF, including the SGS. A simulation of the SGS should therefore include both ORU equipment simulation and some method of including the MIL-STD-1553B bus and its characteristics [3]. Fortunately, PC/XT interface cards are available from several manufacturers, including the ILC Data Device Corporation (DDC) of Bohemia NY. Such cards permit the use of an actual 1553B bus, thereby obviating the need to simulate bus interactions. One would need only interface a simulation running on a PC with an interface card and its associated software to, in effect, simulate an ORU. Several ORU's, running on separate PC's with 1553B cards could then be used to simulate a subsystem like the SGS. An evaluation of the viability of the latter approach proved an aspect of this investigation.

Method

The entire SGS system includes both S-band and Ku-band communication equipment. Given the amount of time available, it was decided to concentrate on the S-band system in general (the so-called Assembly/Contingency Subsystem, ACS, Fig. 1), and the S-band Transponder ORU in particular. It was felt that if a skeleton of this ORU could be simulated along with an interface to the 1553B bus, then greater detail could be fleshed in later (the sensor measurements list is still undergoing changes). A software prototye would be sufficient to establish proof of concept, and this became the main objective of the software development.

Since Ada is the language designated for software development on SSF, it was decided to make its use a requirement of any software simulation. The Ada compilers used was the Alsys DOS Ada Compilation System, Version 4.3.3, and the Alsys FirstAda System. The software drivers [4] supplied with 1553B interface card, DDC model BUS-65517, are written in the C programming language (specifically Microsoft C). This fact required the development of interface packages between the C and Ada languages so that the simulator software could communicate with the 1553B bus. In addition, a user interface to the bus would be needed to communicate control commands to the simulated ORU and to poll it for sensor readings (Control and Monitoring, CM).

Using techniques similar (but not identical) to those described in Ref. 1, it was decided to concentrate first on establishing appropriate Ada data structures to represent the various sensor measurements that were in the latest available requirements for the ACS equipment [5,6,7,8]. Data structures for the sensor measurements of the three major components of the S-band ACS, including the baseband signal processor (BSP), RF Antenna Group (RFG) as well as the transponder were established, and a rough simulation of the transponder was programmed. These various pieces of equipment can be "turned on" and sensor values read. In addition, the transponder can be "controlled" in that internal switches can be set, and signals change dynamically in a rough approximation of what would be expected when this equipment is initialized, and communications established. The control commands are sent to the transponder via the 1553B bus in the form specified in Ref. 6, (including appropriate 1553B subaddressing), and are read from the interface card via the Ada/C software interface packages. For testing and demonstration purposes, a third PC with a different 1553B interface card (SCI Technology, Inc., Huntsville AL, Model BCU/PC-1553) was used as a bus monitor to observe bus traffic (Fig. 2).

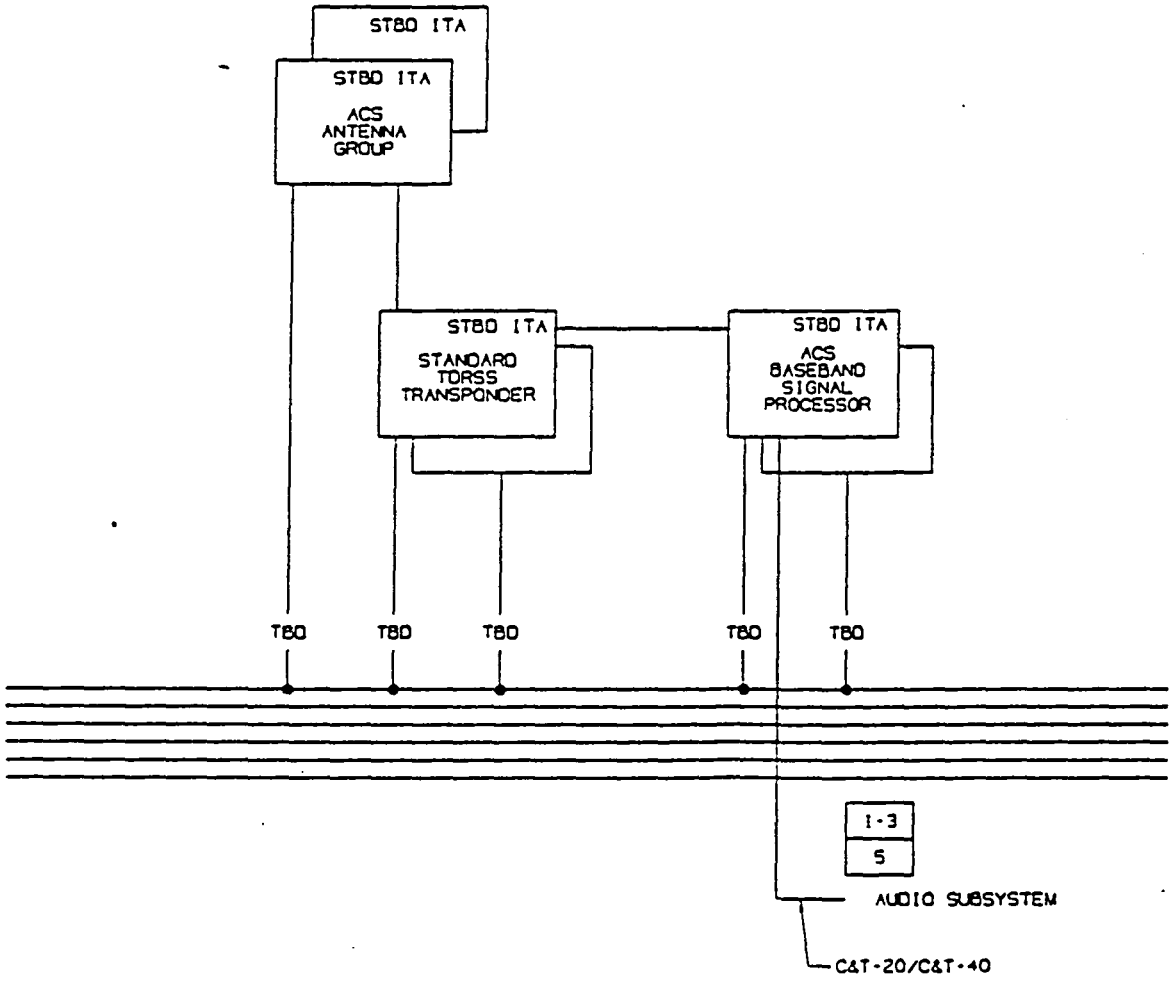


Fig. 1 -- S-Band Assembly/Contingency Space-to-Ground Communications and Tracking Components

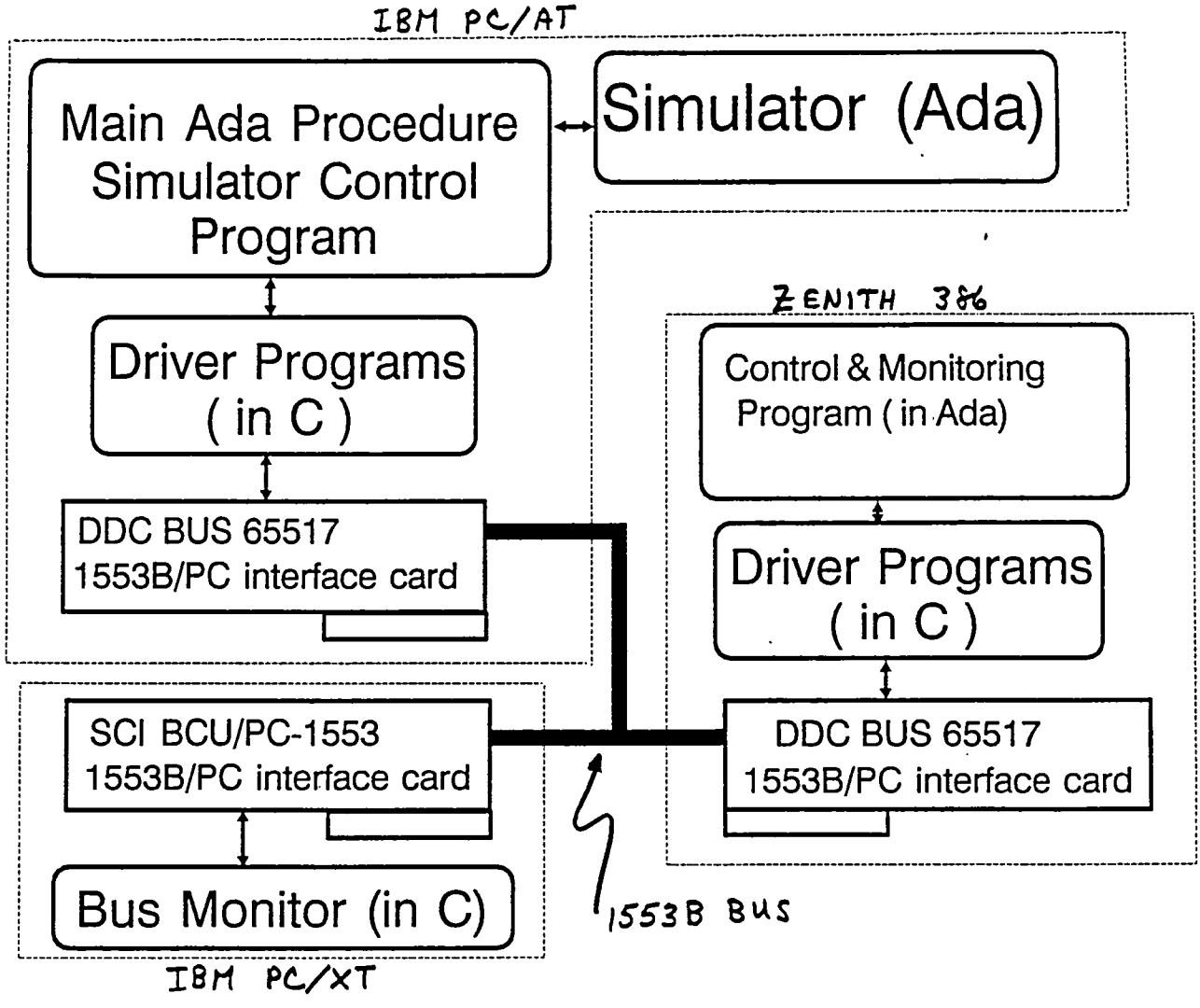


Fig. 2 -- Data flow between software elements and 1553B inter-PC serial connections for simulation testbed

Discussion

A prototype simulation of the S-band Assembly/Contingency Space-to-Ground Subsystem transponder was developed and demonstrated. In addition, a user interface running on a separate PC was developed which permitted Control and Monitoring commands to be sent to the simulator via a MIL-STD-1553B bus and appropriate PC/1553B interface cards. The prototype system indicates that it should be possible to set up a simulation of the space-to-ground components of the Communications and Tracking System of Space Station Freedom using separate PC's. Each PC (with MIL-STD-1553B interface card) could be used to simulate an ORU, and if an actual ORU were available, it should be possible, in principle, to substitute it for the simulation by plugging it into the 1553B bus, and disconnecting the PC simulating it. (Additional "real" signals might have to be applied to the actual ORU to demonstrate signal propagation that was otherwise emulated by the PC's).

The Ada programming language has some features that make it especially suited for this simulation. Chief among them was the "representation clause" construct which made it relatively easy to represent the complex data constructs that are used as Control and Monitoring data. Input/output, however, remains a continuing problem with Ada unless non-standard features are used (packages supplied by various compilers that are peculiar to the compiler). The latter was avoided in this exercise in order to make the code as portable as possible. Much time was therefore spent on the user interface to the system that sends Control commands to the 1553B bus for transmission. (Fortunately, the ANSI standard of "escape sequences" made what would have been an impossible problem at least tractable.)

Acknowledgements

Many thanks to all those who aided in this project. To name some is to omit others-- you know who you are, and I hope that I have sufficiently thanked you all personally.

One person who will be named is my "NASA Colleague", Oron Schmidt. Without his continued support and encouragement, I wouldn't have anyone to thank.

References:

1. "Using Ada Tasks to Simulate Operating Equipment", L. A. DeAcetis, O. Schmidt, and K. Krishen, Comput.Phys.4, 521-525 (1990).

2. "End-to-End Test Capability Test Bed Interface Control Document - Demonstration 3B", NASA/JSC Engineering Directorate, Avionics Systems Division, prepared by Lockheed Engineering and Sciences Company, LESC-25117-3B, Houston TX, May 1989.

3. "MIL-STD-1553 Designer's Guide, Second Edition", ILC Data Device Corporation, Bohemia NY, 1988.

4. "BUS-69008 Microsoft C MIL-STD 1553B Driver, Rel. 1.52", ILC Data Device Corporation, Bohemia NY 11716, July 1989.

5. "Interface Requirment Document, C&TS Local Buses, Preliminary, Volume 11 of 16, CMS-Assembly/Contingency Transmit-Receive Amplifier Interfaces", IRD #10033288, PRELIM-2, GE Aerospace, Camden NJ 08102.

6. "Interface Requirment Document, C&TS Local Buses, Preliminary, Volume 12 of 16, CMS-Assembly/Contingency Standard Transponder Interfaces", IRD #10033288, PRELIM-2, GE Aerospace, Camden NJ 08102.

7. "Interface Requirment Document, C&TS Local Buses, Preliminary, Volume 13 of 16, CMS-Assembly/Contingency Baseband Signal Processor Interfaces", IRD #10033288, PRELIM-2, GE Aerospace, Camden NJ 08102.

8. ACS Transponder Specifications: "B2 Critical Item Development Specification", Specification No. 10033011, Revision B, GE Aeorspace, Camden NJ, March 14, 1991.

AN EXPLORATORY EXERCISE IN TAGUCHI ANALYSIS OF DESIGN PARAMETERS: APPPLICATION TO A SHUTTLE-TO-SPACE STATION AUTOMATED APPROACH CONTROL SYSTEM

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1991

Johnson Space Center

Prepared by:	Don E. Deal
Academic Rank:	Assistant Professor
University & Department:	University of Houston Industrial Engineering Department Houston, Texas 77042
NASA/JSC	
Directorate:	Engineering
Division:	Navigation, Control & Aeronautics
Branch:	Navigation & Guidance Systems
JSC Colleague:	William L. Jackson
Date Submitted:	7 August 1991
Contract Number:	NGT-44-001-800

ABSTRACT

NASA is currently exploring strategies which can contribute to increased efficiency in performance of their various activities and is undertaking to evaluate the potential offered by a multitude of techniques in the Total Quality Management area. One such technique that has received considerable attention within the manufacturing community over the last few years is Taguchi design of experiments and optimization of design parameters.

Taguchi methods are concerned with the quality design of an end-product and give specific focus to efficiency in the design process. The underlying philosophy of the Taguchi approach is that through introducing a planned structure to the experimentation carried out in the formative stages of design that a great deal of information regarding product or system performance can be gleaned in less time and at less cost than through traditional methods.

The chief goals of the summer project have been twofold —— first, for my host group and myself to learn as much of the working details of Taguchi analysis as possible in the time allotted, and, secondly, to apply the methodology to a design problem with the intention of establishing a preliminary set of near-optimal (in the sense of producing a desired response) design parameter values from among a large number of candidate factor combinations.

The selected problem is concerned with determining design factor settings for an automated approach program which is to have the capability of guiding the Shuttle into the docking port of the Space Station under controlled conditions so as to meet and/or optimize certain target criteria. The candidate design parameters under study were glide path (i.e., approach) angle, path intercept and approach gains, and minimum impulse bit mode (a parameter which defines how Shuttle jets shall be fired). Several performance criteria were of concern: terminal relative velocity at the instant the two spacecraft are mated; docking offset; number of shuttle jet firings in certain specified directions (of interest due to possible plume impingement on the Station's solar arrays), and total RCS (a measure of the energy expended in performing the approach/docking maneuver). In the material discussed here, we have focused on a single performance criterion —— total RCS. An analysis of the possibility of employing a multiobjective function composed of a weighted sum of the various individual criteria has been undertaken, but is, at this writing, incomplete.

Results from the Taguchi statistical analysis indicate that only three of the original four posited factors are significant in affecting RCS response. A comparison of model simulation output (via Monte Carlo) with predictions based on estimated factor effects inferred through the Taguchi experiment array data suggested acceptable or close agreement between the two except at the predicted optimum point, where a difference outside a rule-of-thumb bound was observed. We have concluded that there is most likely an interaction effect not provided for in the original orthogonal array selected as the basis for our experimental design. However, we feel that the data indicates that this interaction is a mild one and that inclusion of its effect will not alter the location of the optimum.

INTRODUCTION

A critical stage in the performance of the Shuttle's rendezvous and docking maneuvers with the Space Station will be the final approach, initiated at a separation of a few hundred feet between two spacecraft. This procedure can be supported, either entirely or in large degree, by an automated approach program designed to establish the approach path and control the approach velocity and attitude of the Shuttle relative to the Station, thus guiding the Shuttle into the Station's docking port.

A central goal in the assessment of approach design options is that of determining a superior combination of design parameter settings so as to permit an approach/docking maneuver which, when evaluated against several quantifiable criteria, is projected to produce a near-optimal result. The set of design factors whose optimal levels are to be determined include the glide path approach angle (ϕ), the minimum impulse bit setting (MIB) which defines the mode of Shuttle jet firings, and two gains —— the path intercept gain ($K_{\partial r}$) and approach gain factors (K_x) —— which describe target velocity modification profiles for the approach. The set of performance indicators includes statistics indicative of the terminal conditions at the instant docking is to be achieved (e.g., the terminal velocity and docking offset); fuel expended in the approach (total RCS); and the number of Shuttle firings in specified directions, this last being of concern due to the possibility of plume impingement on the Station solar arrays.

In the summer assignment, we have focused on a single performance criterion, total RCS, which we seek to minimize. Our evaluation medium was a simulation program developed within the Guidance & Prox Ops Section which models the dynamics of the problem and produces estimates for the various performance indicators, all of which are, of course, functions of the design parameter values set by the analyst. The methodology employed in searching for a near-optimum combination of these design factor values was Taguchi parameter optimization, a statistical design of experiments technique, the details of whose analytical steps are given in the sections below.

EXPERIMENTAL DESIGN

The domain for the factor settings in this preliminary study were as follows: Glide Path Angle - $\phi = -10°$, $\phi = 0°$, $\phi = +10°$; Minimum Impulse Bit mode - MIB= norm z, MIB= low z; Path Intercept Gain - $K_{\partial r} = 0.8T$, $K_{\partial r} = 1.0T$, $K_{\partial r} = 1.2T$; Approach Gain - $K_x = 0.8T$, $K_x = 1.0T$, $K_x = 1.2T$, where T for these last two factors is a theoretically derived value posited for each. Thus, our feasible region of interest is defined by four discrete decision variables with three 3-level factors and one 2-level factor.

In an exhaustive search for the optimum set of design factor values, an evaluation of all 54 setting combinations would be performed through a full-factorial experimental design. A by-product of the statistical analysis which follows compilation of the experiment data is a mathematical model which estimates all main factor effects *and* the effects of *all* interactions among these factors on the response variable (here, total RCS). Though

derivation of such a detailed model produces considerable insight into the phenomenon being modeled, there are also obvious disadvantages in this approach. The time and expense involved in running experiments is frequently such that a large number of cases cannot be accommodated.

Fortunately, in the majority of instances, it not necessary to perform a complete enumerative analysis. It has been observed that in engineering design problems the effects of the higher-order interaction terms —— certainly those involving three or more factors —— are very often insignificant and that, in many cases, most of the second-order terms have little effect on the response function. As a result, meaningful information may be gleaned through a considerably smaller battery of experiments, if the experimental design is constructed in a particular manner. These points are, in fact, the premises underlying Taguchi design of experiments.

It not being evident that there would be any significant interactions among the factors, it was decided that we would employ the simplest experimental design that could accommodate all four factors. This was the Taguchi L_9 orthogonal array[1], each of whose four columns could carry a 3-level factor. Note that, as pointed out earlier, one of the design parameters (MIB) was a 2-level factor, the remaining three parameters being 3-level factors. This mismatch between the factor levels of the problem and those provided for in the L_9 array, is only apparent. In actual fact, using a dummy level technique, we simply choose a level designation to repeat as a third level value for the 2-level factor —— for example, MIB_1= norm z, MIB_2= norm z, MIB_3= low z. The net result is that the estimate for the mean effect at one level (MIB= norm z) has twice the precision as that at the other (MIB= low z). In addition, the orthogonality of the experiment array is preserved. The L_9 layout for our matrix of experiments is given in Table 1 below.

TABLE 1: L_9 MATRIX EXPERIMENT LAYOUT

Experiment Number	Factors				Response
	ø	MIB	$K_{\partial r}$	K_x	
1	-10°	norm z	0.8T	0.8T	RCS_1
2	-10°	norm z	1.0T	1.0T	RCS_2
3	-10°	low z	1.2T	1.2T	RCS_3
4	0°	norm z	1.0T	1.2T	RCS_4
5	0°	norm z	1.2T	0.8T	RCS_5
6	0°	low z	0.8T	1.0T	RCS_6
7	+10°	norm z	1.2T	1.0T	RCS_7
8	+10°	norm z	0.8T	1.2T	RCS_8
9	+10°	low z	1.0T	0.8T	RCS_9

[1] A matrix experiment is termed an orthogonal array if the contrasts corresponding to all its columns are pairwise orthogonal (i.e., if the inner product of any two vectors of contrast weightings is zero). See Chapter 3 of Hicks [1] and Appendix A of Phadke [2].

Recall that we wished to evaluate the factor contributions with an eye towards minimizing the total RCS response —— i.e., we were operating from a "smaller the better" basis, as termed in the Taguchi literature. For each of the nine response observations, the *squared deviation* from zero (the "ideal" target in the minimization of a nonzero objective) was computed in accordance with Taguchi's concept of a quadratic quality loss function (see Chapter 2 of Roy [3]). The individual squared deviations were then transformed into dB readings[2] as follows: $R_i = -10 \log_{10}(RCS_i)$. In the transformed domain we then wish to maximize the "signal" response.

An analysis of variance was carried out on the transformed data with the result that all factors were statistically significant, with the exception of $K_{\partial r}$. Figure 1 below displays the estimated factor level effects in the original domain so that we see actual RCS projected effects. The optimum factor settings which are predicted to minimize total RCS are ø = -10°, MIB = norm z and K_x = 1.2T.

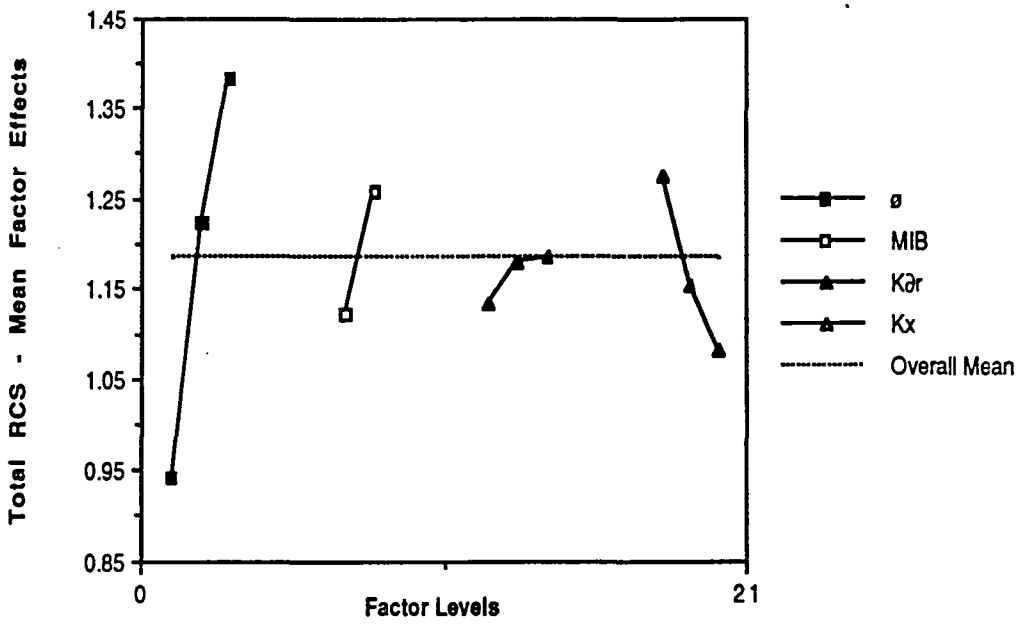


Figure 1: Mean Factor Effects on Total RCS

[2] Transformation of data into decibels is standard procedure in Taguchi analysis and offers an advantage in prediction, particularly when observed values are close to zero but cannot be negative. The transformation concept is also useful in assessing signal-to-noise ratio. See Phadke, Chapter 5.

The sum of squares estimated for $K_{\partial r}$ was pooled with that for the ANOVA error sum of squares to produce an error mean square, or error variance, estimate (this estimate would be used later in comparing predicted values to observed values). The statistical model for the response as a function of the contributing factor settings was then

$$\mu(\varnothing_i, MIB_j, K_{x_k}) = \mu + (\mu_{\varnothing_i} - \mu) + (\mu_{MIB_j} - \mu) + (\mu_{K_{xk}} - \mu)$$

where $\mu(\varnothing_i, MIB_j, K_{x_k})$ is the projected mean response with glide path angle equal $\varnothing_i$, minimum impulse bit mode set to MIB_j and the approach gain equal to K_{x_k};

μ is the overall mean of all observed responses;

$\mu_{\varnothing_i}$ is the estimated effect of $\varnothing$ at setting i;

μ_{MIB_j} is the estimated effect of MIB at setting j;

$\mu_{K_{xk}}$ is the estimated effect of K_x at setting K_{x_k}.

The next step was to compare predicted values to actual observations from the simulation program at various combinations of settings for the contributing factors. Predicted values for the mean response at given factor settings are obtained from the functional description above. Observed means would be estimated (via sample mean values) from several runs made at the corresponding settings through the simulation program, where random initial state errors were included to approximate actual "noise" effects (the nine data points obtained through the experimental design were outcomes resulting from no initial state errors).

Table 2 on the following page lists the model predictions for total RCS at all 18 combinations of the significant design factor levels. Predictions are arranged in ascending order to facilitate a rank-order comparison with observed values in the adjacent column. Observed sample means in the third column are for samples of size 12 at selected level combinations (time did not permit simulation and review of data at all combinations). The standard deviation used in computing the normalized error is a function of the ANOVA estimated error variance and the dispersion observed in the 12 replications at that particular combination of factor levels. We note that the standardized errors are all of reasonable magnitude, except for that at the predicted minimum (factor combination #1), and the rank-order of predicted RCS values agrees closely with that for the observations.

Figure 2, which follows the table, depicts a graphical comparison of the data. The actual range of observed values is about 64% of the range predicted, with the most significant error, again, being that at the optimum combination. However, we note that with this one exception all predictions fall within a 2-sigma error envelope (shown as the area between broken lines in the figure) used as a rule-of-thumb upper bound by some authors. Thus, while the model predictions are not uniformly in close agreement with the corresponding

TABLE 2: COMPARISON OF MODEL PREDICTIONS WITH SIMULATION OUTPUT

Factor Combination Count	Mean - Model Prediction	Mean - Simulation Observations	Error (Percent)	Error (in Std. Dev.'s)
1	.838	.962	-12.88	4.00
2	.895			
3	.940	.997	- 5.68	1.14
4	.990			
5	1.005			
6	1.090	1.112	- 1.96	0.48
7	1.111			
8	1.165	1.180	- 1.26	0.33
9	1.224	1.120	9.29	1.93
10	1.230			
11	1.288			
12	1.307	1.252	4.42	0.94
13	1.314	1.347	- 2.45	0.61
14	1.381			
15	1.445			
16	1.453			
17	1.475	1.369	7.78	1.66
18	1.631			

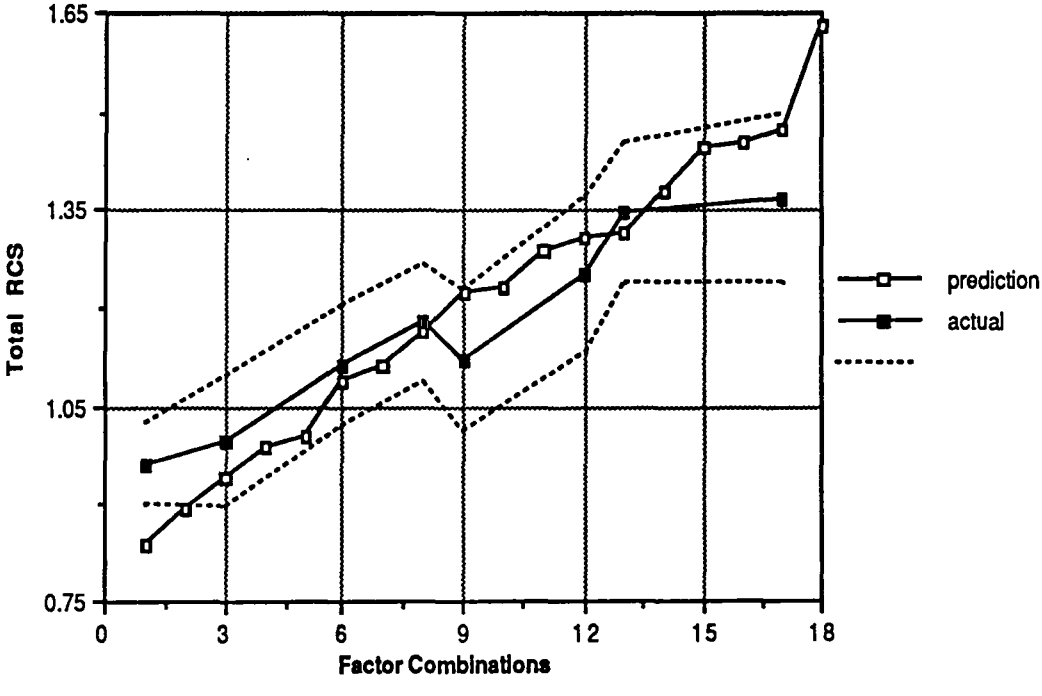


Figure 2: Total RCS Predicted Means vs. Observed Means

observations —— particularly, at the predicted extremes —— the model seems to capture adequately information regarding rank ordering of the possible factor combinations.

CONCLUSIONS

Our assessment is that the primary source of differences between the projections of the statistical model and observations obtained through simulation is most likely a single interaction effect between two of the factors. Recall that the L_9 experimental design employed provided for estimation of no factor interactions. Thus, any interaction effects that are, in fact, present are confounded with some subset of the main effects. However, given the level of agreement between predictions and observations over the range of results, we feel that the postulated interaction effect, though not insignificant, may be a mild one. We emphasize that provision for evaluation of the interaction would likely lead to a mathematical model with improved predictive capabilities; however, our chief objective has been to determine a near-optimum combination of design factor settings, and we feel that we have been able to do that.

For this investigator and the NASA host group, this has been a first exercise with Taguchi methods in a realistic problem setting. Our overall evaluation of the technique is that it can prove to be an efficient analytical tool, particularly once one has gained some experience with the mechanics and develops a feel for its subtleties. As a summary, we express impressions formed and offer some caveats in working with Taguchi analysis.

Rule-of thumb rationale for identifying "significant" factors will not always produce a clear-cut set of contributing variables. In typical ANOVA, a level of significance for statistical tests may be specified, and the F-ratio suggests that either a factor is significant or it is not. This is not the case here. It may be extremely difficult in this application to specify an alpha value in advance and then adhere to the conclusions one is drawn to as a result. A Pareto Principle approach seems more practical. This maxim suggests that approximately half the factors (actually, the factors associated with approximately half the total degrees of freedom) should account for the large majority of the observed variation. These factors are significant, and the sum of squares attributed to the remainder of the variables (i.e., the nonsignificant factors) are pooled with the error variance. This rationale is likely to work well in most instances —— but not always. There will be "borderline" cases where it is difficult to feel confident about assigning a factor to either class. We can only suggest that the best aid in such situations is one's insight into the phenomena under study.

Computation of squared deviations from a target follows from Taguchi's paradigm of the quadratic quality loss function. For cases where deviations from the target are more easily accommodated in one direction than the other, authors in the field offer the useful concept of employing asymmetric penalties wherein computations produce scaled deviations of $k_1\Delta^2$ in one direction and $k_2\Delta^2$ in the other. We feel that in many engineering design problems such asymmetric penalties may be called for.

We have found through our experience with this study that there may be some difficulty, when using small orthogonal arrays with all columns assigned to factors, in determining whether significant interactions are present and in deciding with some confidence which, if any, these might be. The graphical techniques suggested as offering some indication of the presence of interactions are based on comparison of mean effect against mean effect at various levels for two factors. In small experimental designs, these mean effects may be based on a single observation so that the standard error of the mean is relatively large. Evaluation is thus clouded by these uncertainties. A practical method may be to compare predictions against actual observations at various points in the domain as we have done here; lack of close agreement may suggest interaction effects not provided for in the experimental design.

Despite the possibility of overlooking interactions with small, total column-assigned array designs, much information can be gained if we view analysis at this stage as a first step in an iterative process. Frequently, it will be desirable, after performing a preliminary analysis, to "fine tune" the spacing of settings for the significant factors in an attempt to incrementally improve upon the optimum initially identified; thus, a second iteration may be planned as a matter of course. If we conclude that there may exist notable interactions and can determine which these might be, design of the experiment array for the subsequent stage can take this possibility into account.

In evaluating supporting software, we would look for special applications that can assist in selection and modification of the standard orthogonal arrays, advising the analyst as to which experimental design might be most appropriate for the problem at hand. Any software package that performs only the necessary statistical calculations may offer little more than existing nonspecific statistical packages.

Engineering design problems, particularly those undertaken within NASA, typically involve concern with multiple objectives. It may be possible to scale and weight the various individual objectives and pool these so as to enter into the Taguchi analysis with a single (weighted-sum) objective. However, it is not clear whether we can expect results with a similar level of integrity as those deriving from analyses with single-criterion objective functions. It may be necessary to carry out evaluations on a piecemeal basis (i.e., through independent consideration of the various objectives) and then perform a constrained trade-off analysis to ascertain near-optimal solutions.

We should underline the fact that the philosophy of Taguchi design optimization is to produce a quality design through a structured experimentation process which yields a maximum amount of information about the system under study with a minimum investment of time, effort and cost. As with many other methods within the TQM field, realization of increased efficiency depends on a shift of major effort to the early steps in the project. In Taguchi analysis it is *engineering experience and intuition* that are called upon at the beginning of the process in postulating controllable design factors and potential interactions and in establishing a feasible region of parameter settings to produce a desired result. Failure to take advantage of these elements can reduce the return promised by the methodology.

BIBLIOGRAPHY

[1] Hicks, C.R.; *Fundamental Concepts in the Design of Experiments* ; CBS College Publishing, 383 Madison Avenue, New York, N.Y. 10017; 1982

[2] Phadke, M.S.; *Quality Engineering Using Robust Design* ; Prentice Hall, Englewood Cliffs, N.J. 07632; 1989

[3] Roy, R.K.; *A Primer on the Taguchi Method* ; Van Nostrand Reinhold, 115 Fifth Avenue, New York, N.Y. 10003; 1990

LUNAR PRODUCTION AND APPLICATION OF SOLAR CELLS, AND SYNTHESIS OF DIAMOND FILM

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1991

Johnson Space Center

Prepared by:	P. H. Fang, Ph.D.
Academic Rank:	Senior Scientist
University & Department:	Institute for Space Research Boston College Chestnut Hill, Massachusetts 02167
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Solar System Exploration
Branch:	Mission Science and Technology
JSC Colleague:	Thomas A. Sullivan, Ph.D.
Date Submitted:	August 9, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

This report will describe two projects which are carried out under the Summer Faculty Fellowship Program-1991:

I. A conceptual design of a solar cell manufacturing plant on a lunar base. This is a large program that requires a continuous and expanded effort and the present report will reflect a current status.

II. An experiment on the synthesis of diamond film. Encouraging, but not yet conclusive evidence has been obtained on a new method to synthesize diamond film. The procedures and observations are presented in this report.

A third project is an analysis of the solar cell performance over five years on the moon based on Apollo missions. A paper has been completed and will be submitted to the journal Solar Cells for publication. The paper is co-authored with J. R. Bates of JSC who carried out the experiment and collected the data about twenty years ago. Due to the page number constraint of this report, this paper will not be included but will be available separately.

Following is a table of contents:

A Solar Cell Production Plant at a Lunar Base

Introduction

Electric power is a requirement for lunar exploration. One method to produce this power is through a conversion of solar energy by a photo-voltaic device known as a solar cell. A good example is the solar cell panel of the seismic experiment of the 1969 Apollo 11 mission. The total power level of these panels was about 40 watts. Since that time, there has been much progress in solar cell technology and larger scale solar cell panels have been realized. Multi-kilowatt solar panels have powered satellites and several megawatt power solar panels have been constructed for terrestrial operations.

Through the analysis of rocks and soil retrieved from the moon, it has been recognized that the materials required for fabrication of solar cells can be provided by materials indigenous to the moon[1]. A solar cell production process based on vacuum and heat, both of which are naturally provided by the lunar ambient can be envisioned. Thus in addition to an indigenous material utilization, an indigenous environment utilization can be realized. This approach would greatly reduce the reliance on Earth resources which would have to be transported to the lunar site.

1) Solar Cell Components and Processing

The solar cell adapted in the present development is based on a thin silicon film. As a thin film, a support structure in the form of substrate is required. The physical aspect of the silicon component and the substrate has been discussed previously[2]. The present report will be devoted to the production aspect.

(i) Substrate.--There are various kinds of substrate which can be used to make solar cells. They are listed in the following table:

Substrate	Lunar source	Fragility	Production Facility	Production Mode
silicate glass	Available	Fragile	Complex	batch process
iron, titanium or steel	Available	Stable	Complex	continuous process
polyimide	Not available	Stable	Complex	continuous process
metallic glass	Available	Stable	Simple	continuous process

In terrestrial application, silicate glass, steel and polyimide have been used. For the lunar case, other materials listed above can also be considered due to the absence of corrosion elements. One deciding factor on the choice of the substrate would be the lunar availability. In this case, polyimide would not be practical because of the lack of starting material. Another point of consideration is based on a production schedule: If solar cells are to be made before a capability of lunar industry to produce the substrate exists, and the substrate has to be transported from the earth, glass would not be a practical choice both because of its fragility and weight. For this analysis, suffice that the solar cell fabrication is not greatly dependent on the specific types of the substrate.

(ii) Silicon Layers.--The semiconductor silicon layers, either in amorphous or microcrystalline form, are to be deposited from the vapor of molten-silicon heated in a vacuum. The preparation of the silicon layers has been described previously[2] , the production procedure will be described under the production plan

(iii) Hydrogenation.--Thin film silicon solar cells, including amorphous silicon and microcrystal silicon, require hydrogen passivation of the inherent defects. In the conventional chemical vapor deposition, the starting silicon material is a hydride and contains rich

hydrogen component. Thus the hydrogenation is accomplished in situ during the silicon deposition. The power output of solar cells made by this process shows a deterioration of about one third after the first month of operation under solar illumination and a somewhat slower decay afterwards. This instability is explained by a formation of polymers of SiH_2 due to an excessive hydrogen content in the deposition process[3]. The formation of this polymer is greatly reduced in a two step process adapted in the present approach of a deposition of silicon first by evaporation and followed by a post hydrogenation. The hydrogenation involves annealing the deposited material in a hydrogen plasma at about 300° C. In the current laboratory practice, hydrogen flows continuously at a pressure of approximately of 1 torr. At a lunar base, hydrogen will be scarce and conservation is highly desirable. A method to achieve this purpose is to replace the continuous hydrogen flow with a stagnant operation where the system is closed with only a limited hydrogen replenishment to compensate for the hydrogen diffused into the silicon body. Such an operation is not possible in the chemical vapor deposition system since, in that case, besides the silicon hydride, an amount of doping impurity is removed. The relative concentration, which should be maintained at a constant value, would vary due to the different removal rate of the dopant and of silicon. However, in the present approach, there is only a single hydrogen component and thus a stagnant operation becomes possible. A laboratory experiment to confirm this approach would enhance the practicality and the economy of the lunar production process.

(iv) Electrode Formation.--Depending on the detailed solar cell structure, the electrode material would be different. In the case of a combination of a top microcrystal layer on the amorphous layer, due to an adequate electrical conductivity of the microcrystal silicon, a grid electrode configuration can be applied. Two electrode materials which would be suitable for lunar solar cells are aluminum and titanium. These materials are easily oxidized and become

electrically resistive in the terrestrial environment. The oxidation problem does not occur under the lunar vacuum condition.

A different case is that when the solar cell is made completely of amorphous silicon, a large area coverage instead of discrete grid lines becomes necessary. In this case, the electrode material has to be both optically transparent and electrically conductive. Oxides of several metals such as indium, tin or zinc satisfy these requirements, but these metals are rare in the lunar rocks.

The electrodes described above are for the top electrode where the light enters. The bottom electrode can be the substrate itself when the substrate is conductive. In this case, a solar cell with a large area would provide large electrical current but the voltage would be limited to about one volt. To produce a high voltage, the large area has to be dissected and reconnected in series. A means to make this connection is by two pairs of of terminal clamps such that the terminals of one pair is connected to the top of one cell and the bottom of the second cell, and vice versa for the second pair. The clamps will be designed such that the connections can be made mechanically and robotically.

A different situation is that when the substrate is not conductive. In this case, solar cells can be a continuous large sheet and the electrode patterns, with a combination of series and parallel connections can be deposited on the substrate as well as on the top surface of the solar cell.

2) Production Planning

Two schemes of the production plant can be considered, a full production plant and a smaller pilot production plant. The full production plant, with an annual solar cell production rate of multi-megawatt powers can be realized based on an extrapolation from ensuing discussions. The present report will devote to

a pilot plant in order to acquire a operational experience and facilitate an earlier realization of the lunar solar cell production.

Processing chambers of this pilot plant will consist of three independent modules:

i) Silicon deposition chamber.--There are three or possibly four different silicon layers. These layers will be processed time sequentially in a single module. Inside the chamber, two different silicon crucibles will be placed, one for intrinsic silicon and one for p-type silicon alloyed with boron. N-type silicon can be made by a coevaporation of intrinsic silicon and antimony. In operation, a stepping mechanism will place the crucible with the different silicon types into the focal point of the concentrated light beam. In between successive depositions of different silicon layers a louver will be manipulated to block off the incoming light beam such that a stray deposition does not occur.

Finally, for the microcrystal silicon layer, the substrate temperature has to be modulated. One method to achieve this temperature is to adjust the louver of the substrate heater window to provide an elevated temperature of 550°C. This temperature would preclude many substrate materials. A lower temperature growth of microcrystals has been reported through a reactive chemical vapor deposition[6]. Whether this can also be realized in vacuum evaporation is yet to be investigated.

In order to reduce the large range of the tracking angle of the solar concentrators, which in practice would require multiple optical elements, restriction of the operation time would be helpful. In this case, instead of a whole 340 hours of lunar day, only a fraction of a lunar day will be operative. Take this fraction to be 2/3, the time equivalent will be 227 hours and, the total production rate would be reduced. For a solar cell area of $0.1m^2$ and a complete deposition time, including changing the silicon source, to be 0.5 hour, the total area of solar cells produced per a lunar day will be $0.1m^2$ x 227 hours/0.5 hour= $45.4\ m^2$. Assuming

conservatively a solar cell conversion efficiency of 5%, the power output becomes 45.4m^2 x 70 w/m^2 = 3.2 KW. For an earth year, the total solar cells will produce 38 KW. This number almost coincides with the solar panel output of the Freedom Space Station under present design of 37 KW.

A distinct feature of this chamber is an absence of high voltage and high power leads. These leads are numerous in the common vacuum deposition chamber and the necessary electrical isolation takes up much space in the chamber. Our conceptual vacuum chamber would have a much simpler interior and the space utilization would be more efficient. However, there is also a problem which is less important in the ordinary system in which the window is merely a view port there. In our system, the window becomes the entrance port of the light beam as a power source and high transparency is of utmost importance. A technique to maintain a clean window would be an important subject.

b) Hydrogenation chamber.--While in the silicon deposition chamber, the substrate should be in the line of sight of the silicon source; therefore, multiple solar cells can not be stacked for simultaneous deposition. In the hydrogenation chamber, however, the solar cells can be stacked with a separation of about the mean free path of the plasma ions. At 1 torr pressure, the mean free path is on the order of 1 cm. The required temperature is about 300°C which will be obtained through solar heating.

The hydrogenation process requires about one hour [2], which is longer than the silicon deposition time. However, the output from silicon deposition chamber can be accumulated to batches, and the production can thus be streamlined.

c) Electrode attachment chamber.--This chamber will deposit electrode materials by a vacuum evaporation by solar heating with a masked grid or other desirable electrode configuration and the adapted metal or oxide materials.

3)Space System Engineering

The system involvement can be divided into several components:

(i) Material sources.--The two chief materials are silicon and metallic or glass substrate. Silicon extraction from lunar rock has been investigated in several contract reports[7,8]. Further purification during the vacuum deposition can be made with the same principle of a fractional distillation.

At present, there is no specific effort on substrate construction. This can be carried out when the lunar solar cell production project is dedicated and the size of funding is adequately scaled.

(ii) Heater system.-- The basic component of the solar heater system are a set of optical lenses and mirrors. The present report only studies the general concept. Further effort with expertise is evidently necessary in order to build a real system.

(iii) Machinery.--Under this category is the control and monitoring system, the motor drive to manage the transfer of the substrate, the optical tracking system, and the interchange of the deposition materials. These systems, including robotic features, already exist in terrestrial industry and therefore an adaption of these systems will be basically a development process.

iv) Physical size, weight, and cost analysis --The physical size of the manufacturing plant, including components described in iii) above can be packed in a size of about 1 m^3, and the weight would be sub-tons for a pilot plant. For production plant, the value would be less than one order magnitude higher.

For the cost, in earth industry, the chief cost is the plant facility and labor which represents 70%[9]. In the lunar case, these aspects are not applicable and the only direct cost would be that under iii). Cost of such systems may be in the range of several hundred thousand dollars when applied on earth. When these systems are to be employed on the moon, the

packaging and the reliability requirement could increase the cost by perhaps one order of magnitude.

4) Conclusion and Recommendation

I came to Johnson Space Center under the Summer Faculty Fellowship program with a background of solar cell technology and some planning to produce solar cells on the lunar base. Through the eleven weeks of concentrated study and intensive interaction with the scientists in this Center, in an environment where one sees the giant rocket on the entrance of the Center, pictures of the lunar surface and the astronauts on the hallway, and the real lunar rocks are almost adjacent to my office, my concepts are clarified and my conviction is deepened that the lunar base is a fertile ground to support a solar cell manufacturing: There are materials on the base which , after some processing, can be the components of the solar cell. The vacuum as a lunar ambient is a built-in resource for solar cell processing. The quality of the vacuum the earth laboratory can only envy. The Sun as a power source for solar cell deposition is available only half of the time, that is, during the lunar day, albeit cloudless and stable.

The basic idea of photovoltaic operation to convert solar energy into electricity on lunar base is as applicable as it is on Earth. However, there are some differences. One example is amplified through my recent work at JSC on the analysis of solar cell output data from several Apollo missions: due to the prolonged heating period with heat removal only through an inefficient black-body radiation, the solar cell temperature would be very high and the solar cell performance would be highly inefficient if one transfered directly the terrestrial solar cell structure to the lunar solar cells. An emphasis has to be paid on the emissivity and reflectivity of the solar cell even more stricter than the space solar cell-the heating period of space solar cells is much shorter than that of lunar cells.

The concept of solar cell processing I have developed has a built in simplicity that makes the process practical in necessarily restrictive lunar conditions. However, there are

numerous detailed steps to be experimented and established, some of these have been discussed in the body of the report. In consideration of technological and engineering effort involved in this project, it is high time to invest a concerted effort for this purpose.

I recommend that an implementation of a lunar solar cell manufacturing project should be carried out, involving both laboratory work and system planning, to correct the current unsatisfactory status of meager efforts of a few people. At present, opinions can be found in information media and the technology community as the positive encouragement on lunar and martian exploration, and solar cell manufacturing has been enumerated in a JSC document: Using Space Resources[1] and cumulated in Architecture IV of America at the Threshold[10]. I propose that steps should be taken commensurrate with these two august documents.

References

1. T.A. Sullivan and D.S. McKay, Using Space Resources, Mission Science and Technology Office, Solar System Exploration Division, NASA Johnson Space Center(1991).

2. P.H. Fang, C.C. Schubert, Peiguang Bai and J.H. Kinnier, Appl. Phys. 41, 361(1982).

3. W.A. Nevin et al. Appl. Phys. Lett. 58,2669(1981).

4. P.H. Fang and Ruguang Chen (unpublished)

5. P.H. Fang, Solar Cells, 25, 31 (1988).

6. P.H. Fang, Properties of Amorphous Silicon (2nd Edition) emis Data Rev. Ser. (INSPEC Publ. London, 1989) Sec. 8.19, p. 306 for reference therien.

7. W.N. Agosto and Su Jin Yun, Separation of Metal Oxides in Industrial Fly Ash by Fluoroacid and Ion Exchange Technique and its Application to In-situ Lunar Resources Utilization, Contract NASA/JSC T-9720P from Lunar Industries Inc. December.

8. R. Keller, Manufacture of High Quality Silicon, an interim report of Dry Extraction of Silicon and Aluminum from Lunar Ores, Contract NAS9-17811 from EMEC Consultants, August. 18, 1989.

9. R.A. Whisnant et al. Eighteenth Photovoltaic Specialist Stafford Conf. 1985 p.1537-1544.

10. T. P. Stafford, Report of the Synthesis Group on America at the Threshold, May, 1991

II. An Experiment on a Diamond Growth Method

In recent years, there has been a proliferation of activities to grow diamond, based on the decomposition of a gaseous hydrocarbon source as suggested thirty five years ago by B.V. Spitsyn and B.V. Derjaguin (USSR Inv. Cert. no. 339134, 1956). One deficiency of this method is that the diamond grown by this approach is confined to polycrystals of micron sizes. Crystals of millimeter or larger have been grown so far only on natural diamond faces by homo-epitaxy. A new method which potentially can grow large single crystals by carbon ion implantation on metals has been reported by Prins and Gaigher recently (Second International New Diamond Sci. and Tech. Conf. Sept.,1990 Crystal City, Va.). These metals should have a lattice matching to diamond and have a low carbon solubility. To overcome the disadvantage of using ion implantation as a means of introducing carbon, I have been working on a different approach based on a controlled diffusion of carbon through the same type of metals of Prins and Gaigher. In this method, a paste of an organic suspension of fine grain carbon black is painted on one side of the metal foil surface. By heating in a vacuum or in a protective atmosphere, carbon diffuses through the foil and resurfaces on the opposite surface. The carbon atoms, when they have sufficient mobility, could grow into diamond through a hetero-epitaxy with the host metal.

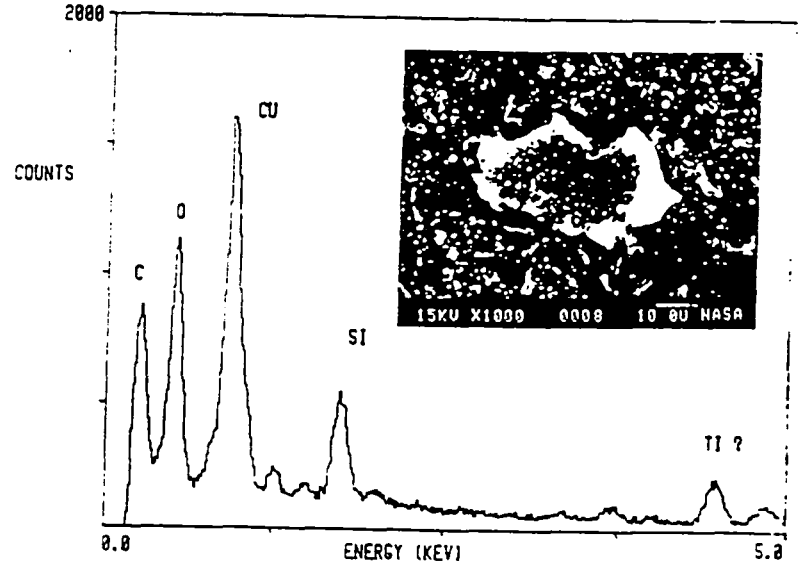
In my work prior to arrival at JSC, I have used vacuum ambient and some preliminary experiments have been made to determine desirable parameters, including the temperature and time window. After my arrival at JSC, I have found that there is a convenient facility in the form of a tube furnace with a controllable protective gas. In my previous vacuum arrangement, the metal in the form of foil strips was resistively heated by passing electrical current through the metal. The resulting temperature is highly non-uniform from near the electrodes to the mid-section. Furthermore, the temperature was measured by an optical

pyrometer. In the low temperature range used in the experiment, around 600 to 1000°C, there is a temperature uncertainty of around 50°C. With the use of a tube furnace, uniformity and accuracy evidently would be greatly improved.

In my previous arrangement for the protective gas, I used low pressure hydrogen in combination with a mechanical pump. At JSC, since there is a safety concern and a hydrogen removal system cannot be readily made, I have used argon at slightly above one atmosphere pressure (suggested by Amy Jurwitz.). In this case, the excess gas is passed through an oil bath to prevent an air backflow and the mechanical pump is eliminated.

I have carried out the experiment with some encouraging results:

1. The carbon diffusion through the metal foil and outdiffusion to the opposite foil surface has been confirmed by an energy dispersive spectrum analysis (EDX) shown in the following figure in combination with a scanning electron microscope (SEM) picture in the



inset. This speciman is made with the following conditions: The metal is a copper foil of 25 micrometer in thickness. After a purge with argon of 17 lbs pressure, namely, about three llbs. above the atmosphere pressure, for two hours, the temperature is raised to 925°C in half an hour and maintained at this temperature for six hours. After that heater power is turned off and the speciman is naturally cooled to about 60°C in three hours.

The EDX picture shows that, besides carbon and the substrate copper, there is a considerable amount of oxygen, silicon and possibly also titanium in the general area excluding the large pattern in the center of the inset SEM picture. While the large pattern is completely carbon according to EDX, the small bright specks are oxides of various metals and of silicon.

The pressure of oxygen is evidently undesirable as it would etch away the carbon, including the diamond crystal. This etching effect might explain an isolated occurrence of the large carbon patterns. An effective method to eliminate the oxygen would be to mix a small amount of hydrogen into the argon gas. This approach, suggested by G. Lofgren, can be carried out easily by pre mixing of the two gases, but the available time prevents this attempt and has to be awaited for a later time. Future effort will be directed towards the crystallographic and Raman scattering identification of the carbon specks such as the large central pattern in the inset. Once the diamond structure has been proved, the next objective would be a realization of an epitaxial growth such that the bright specks become a large area film coverage over the whole grain of the metal, eventually to a large single crystal. This plus a formation of the oxides is the original concept of this work.

EVALUATION OF FUNGAL METABOLIC COMPOUNDS RELEASED TO THE AIR IN A RESTRICTED ENVIRONMENT

Final Report
NASA/ASEE Summer faculty Fellowship Program - 1991
Johnson Space Center

Prepared By:	Robert N. Ferebee, Ph.D.
Academic Rank:	Associate Professor
University and Department:	University of Houston - Clear Lake Department of Biological and Allied Health Sciences Houston, Texas 77058
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Research and Operations
JSC Colleague:	Duane L. Pierson, Ph.D.
Date Submitted:	August 20, 1991
Contact Number:	NGT-44-001-800

ABSTRACT

The metabolic action of selected fungi species on common components of the interior of Space Station Freedom (SSF) will be tested. When present volatile organic chemicals will be collected on porous polymer adsorbent columns. Using thermal desorption the volatile compounds will be passed onto a gas chromatographic column for analysis.

The Space Station Freedom (SSF) modular complex will largely be individually self contained and the established air environment will not be easily adjusted. The development and maintenance of a safe working environment offers a considerable challenge. Present plans for use of SSF acknowledge periods of manned activities and alternate times when the station is unmanned. The obvious necessity for clean and safe air and water during periods of use have been pursued as fundamental systems to SSF success. Somewhat less obvious although perhaps of no less importance to the success of long term cyclic usage are those periods of inactivity. It is during these periods when spores from microorganisms may be afforded the best conditions to germinate and in the vegetative form react with the complex synthetic chemical polymers which compose the furnishings and hardware of SSF nodes.

Biodegradation could constitute a real hygiene problem , if the organisms form and release volatile organic chemicals. Similar problems have been documented in closed and improperly ventilated buildings and work spaces. Many of the metabolic products of fungi and bacterial growth create a variety of health problems.

Analytical chemical techniques will first be used to document the growth of Aspergillus, Penicillium,and Cladosporium fungal species on the potential substrates Nomex and Kevlar. Any volatile organics that are released will be measured using the spectrum of gas adsorption chromatography.

The level of microbial contamination that is necessary to produce such volatile compounds and the relative amounts expected to accumulate will be estimated.

INTRODUCTION

Air Quality is among the factors that will limit the ability of space explorers to participate in the extended operations of outer space.A great part of the concern is not with the obvious wastes from respiration, which is being dealt with effectively, but rather with the more discrete space travelers. The microorganisms upon being placed in environmental conditions that favor growth can sporulate, germinate, etc. and during subsequent metabolic processes release volatile organic compounds. Such volatile organics have been documented regularly and to have varying degrees of toxicity(3). There is a growing file of data for release of volatile products from the growth of both molds and bacteria on food stuffs(1,2). Food grains, such as barley, corn, oats, and wheat have all been used as substrates for growth of a variety of fungi. Organic compounds from the fungal growth include several alcohols and terpenes from species of Aspergillus, Fusarium, and Penicillium(1). Most of these volatile organics will pose no more than minimal risk, while other compounds including the terpenes are referenced as being moderately toxic when ingested or inhaled. Space Station Freedom(SSF) represents a largely closed environment that will require the treatment and recycling of its air. The growth and metabolic activity of groups of organisms such as fungi may add to the air quality considerations.

PROBLEM

This study was undertaken to address one of the air quality concerns for SSF. Three fungal genera were planted on media representing both natural and synthetic compounds. After allowing time for growth of the fungi the gas from the culture vessel headspace was tested for volatile organic chemicals.

EXPERIMENTAL DESIGN

Fungal cultures were obtained from the microbiology section of the Biomedical Research and Operations Branch at NASA/JSC. The particular fungi were chosen as representative of those previously isolated from surface swabs on Skylab and Space Shuttle missions(5) and because they have all been cited as producing volatile metabolites from a variety of substrates(2,7). A culture battery consisting of six 250 ml, 70mm wide-mouthed glass containers fitted with teflon lined lids and stainless steel luer lock bulkhead ports were used to contain the fungal cultures (Fig. 1).

The glass containers housed the media filled 54 mm glass petri plate cultures of the selected fungi. Each culture vessel was fitted with 7mm glass tubing and glass wool spore traps for both affluent and effluent air. The affluent air was humidified by passage of compressed air through a glass trap filled with sterile deionized water. The affluent air was further treated by passage through a tenax-gc 60/80 mesh organic adsorbent filter (Analabs Dist. for Chrom. Tech. Inc., Houston, Tx.). Collection of the headspace metabolites was by passing a measured volume(10ml/min.+/- 1.0 ml), filtered and humidified air over each of the cultures from a single air source. Each culture vessel was sampled continuously for a period of 12 days. The effluent air was passed through 6 mm x 2mm x 11.5 cm multi-bed carbon adsorbent glass filter tubes (Carbotrap 300 multi-bed thermal desorption tube, Supelco, Bellefornte,P.A.). A total of 144 L of air was sampled for each of the cultures and controls.

IDENTIFICATIONS

Effluent air adsorbed on the carbotrap tubes from six cultures and three controls were thermally desorbed and passed through the column of a gas chromatograph that is fitted with a mass selective detector (Hewlett-Packard Series II 5890 GC and HP 5971A MS, Sci. Div.,Palo Alto,CA.). A fused silica capillary column 60M x 0.25 mm I.D., film thickness 0.12 um (SPb tm-1, Supelco,Inc.) was used for resolution of the organic compounds. Helium gas of high purity was used as the carrier, the sample volume was 0.6+/- 0.1 ul and the injection temperature set at 200 C. The desorption products were cold trapped for three minutes in the first 1M of the column. Retention times and compounds from the first experiment are given in (Table 1).

DISCUSSION

All of the fungi exhibited good growth in the experimental environment chosen as evidenced by increase in mycelial mass and microscopic observation of the hyphae. The penicillium was the only fungi that evidenced any significant release of organic compounds. Traces of 2-butene, n-propanol, and octanol were found from the sampling of Penicillium pupurogenum growing on malt extract agar(MEA). Additionally a significant quantity of ethanone - 1 phenyl was released from the penicillium culture on MEA plus Nomex fabric; the GC-MS scan for this is included (see appendix). Additional analysis will be needed to support this finding. Aspergillus flavus exhibited traces of ethanol and 2-propanol. The Cladosporium species gave trace amounts of octanal but not on MEA plus Nomex. All of the fungi used exhibited hardy growth characteristics. Additional experiments with the Penicillium pupurogenum and with other penicillium species are planned for the fall of 1991 and spring of 1992.

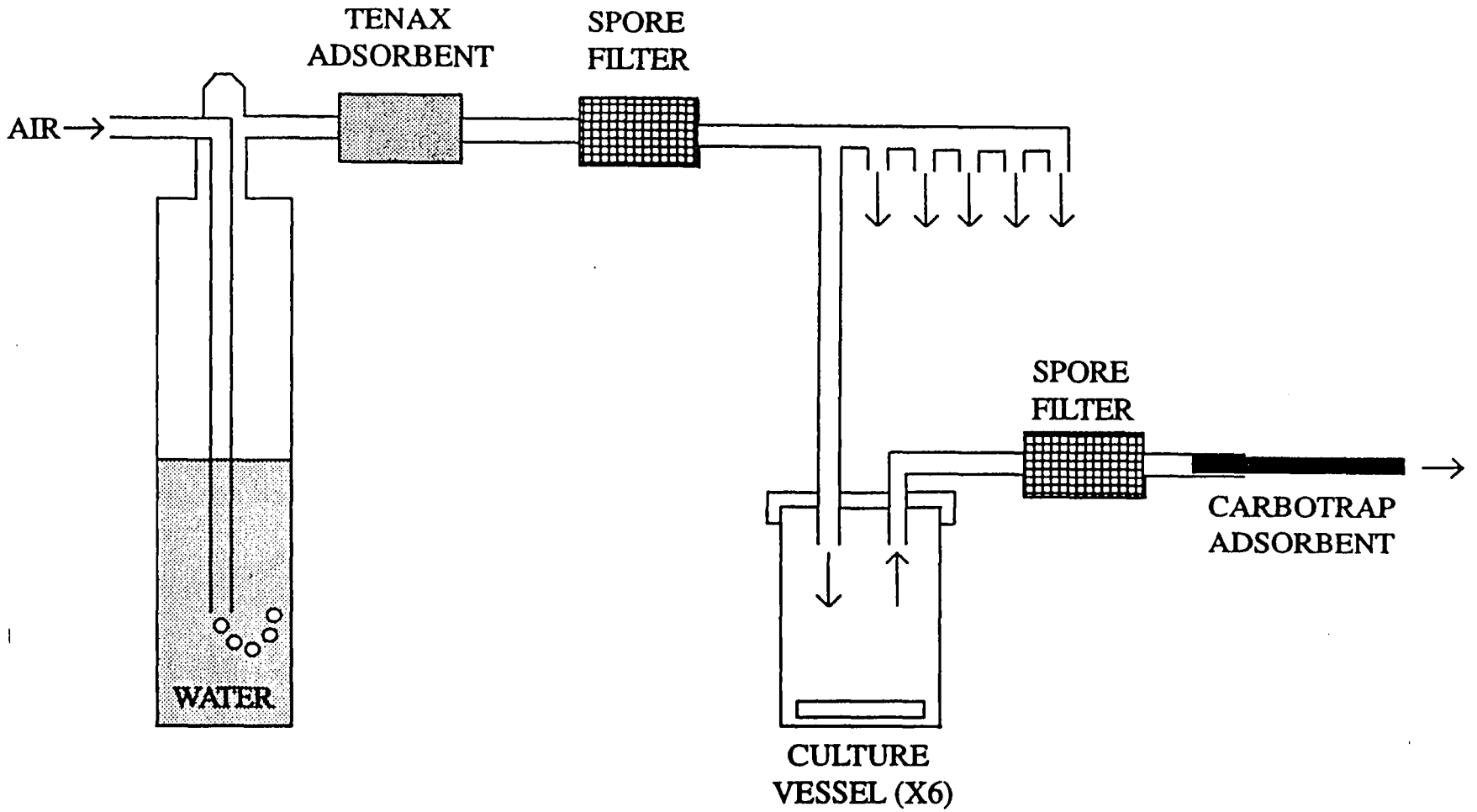


FIG. 1 - CULTURE SYSTEM FOR COLLECTING VOLATILE ORGANICS

Table 1. Volatile Metabolites - 12 Day Cultures of Fungi

Organism/Std	Retention Time	Compound	Reliability%
Standard	7.40	Ethanol	72
"	8.80	Tert-Butyl Alcohol	-
"	10.10	N - Propanol	-
"	16.66	1 - Octanol	-
"	18.00	D - Limonene	83
"	20.60	α - Pinene	91
"	21.67	ß - Pinene	91
Penicillium purpurogenum	9.30	2 - Butene	80
"	9.92	N - Propanol	-
Penicillium + Nomex	11.18	Ethanone 1-phenyl	95
Penicillium	25.90	Octanol	96
Aspergillus flavus	7.60	Ethanol	64
"	12.00	2 - Propanol	72
Aspergillus + Nomex	_	No compounds	-
Cladosporium sp.	27.4	Octanal	87
Cladosporium sp. + Nomex	_	No compounds	_
MEA Control	-	No compounds	-
NOMEX Control	-	No compounds	-

Literature Cited

1. Borjesson,T., U Stollman, P. Adamek, and A. Kaspersson. 1989. Analysis of volatile compounds for detection of molds in stored cereals. Cereal Chem. 66: 300 - 304.

2. Borjesson,T., U. Stollman, and J. Schurer. 1990. Volatile Metabolites and Other Indicators of Penicillium aurantiogriseum Growth on Different Substrates. Appl. Environ. Microbiol. 56(12): 3705 - 3710.

3. Budinski,Kenneth G.1983. Engineering Materials, Properties and Selection.2nd Ed. Reston Pub. Co., Reston, VA.

4. Matcham, S.E. , B. R. Jordan, and D. A. Wood. 1985. Estimation of fungal biomass in a solid substrate by three independent methods. Appl. Microbiol. Biotechnol. 21: 108 - 112.

5. Pierson, D.L. and N.M. Cintron. 1985. Microbiological Report in Support of Spacelab 3 Mission. NASA/ JSC-22202. 1-42.

6. Seitz, L. M. , D. B. Sauer, R. Burroughs, M. E. Mohr, and J. D. Hubbard. 1979.
Ergosterol as a measure of fungal growth. Phytopathology. 69: 1202 - 1203.

7. Sprecher, E., and H.P. Hanssen. 1982. Influence of strain specificity and culture conditions on terpene production by fungi. Planta Med. 44: 41 -42.

8. Wilkins, C. K. , and S. Scholl. 1989. Volatile metabolites of some barley storage molds. Int. Jour. Food Microbiol. 8: 11 - 17.

TARGET DETECTION USING FRACTAL GEOMETRY

Final Report

NASA/ASEE Summer Faculty Fellowship Program

Johnson Space Center

Prepared by:	J. Joseph Fuller
Academic Rank:	Assistant Professor
University & Department	WV Institute of Technology Department of Math/CS Montgomery, WV 25136
NASA JSC Directorate	Information Systems
Division	Information Technology
Branch	Software Technology
JSC Colleague	Timothy F. Cleghorn, Ph.D.
Date Submitted	August 2, 1991
Contract Number	NGT-44-001-800

ABSTRACT

The concepts and theory of Fractal Geometry were applied to the problem of segmenting a 256 x 256 pixel image so that manmade objects could be extracted from natural backgrounds. The two most important measurements necessary to extract these manmade objects were fractal dimension and lacunarity. Provision was made to pass the manmade portion to a lookup table for subsequent identification.

A computer program was written to construct cloud backgrounds of fractal dimensions which were allowed to vary between 2.2 and 2.8. Images of 3 model space targets were combined with these backgrounds to provide a data set for testing the validity of the approach.

Once the data set was constructed, computer programs were written to extract estimates of the fractal dimension and lacunarity on 4 x 4 pixel subsets of the image. It was shown that for clouds of fractal dimension 2.7 or less, appropriate threshholding on fractal dimension and lacunarity yielded a 64 x 64 edge-detected image with all or most of the cloud background removed. These images were enhanced by an erosion and a dilation to provide the final image passed to the lookup table.

While the ultimate goal was to pass the final image to a neural network for identification, this work shows the applicability of fractal geometry to the problems of image segmentation, edge detection and separating a target of interest from a natural background.

INTRODUCTION AND MOTIVATION

Much of the work in automatic imaging of space scenes has been done assuming a dark sky background. This scenario, however, is the exception rather than the rule. Quite often a target will have the earth as its background thus providing a much more difficult problem for automatic image segmentation.

It has been observed that natural scenes are usually fractal in nature and have distinctly different fractal characteristics from those of manmade objects. Thus it should be possible to use Fractal Analysis as a way to separate manmade objects from a natural background. In this paper, we explore ways of doing this separation in a timely and efficient manner.

NECESSARY CONCEPTS FROM FRACTAL GEOMETRY

Fractal Dimension - Fractal Geometry is a branch of mathematics developed by Benoit Mandelbrot and detailed in the book The Fractal Geometry of Nature [3]. In this work, Mandelbrot defines a new measure of the dimension of a set called the fractal or similarity dimension. This dimension is based on the notions of similarity and self-similarity.

A set A is said to be self-similar if it can be divided into N disjoint subsets each of which are exact replicas of the original set. A is said to be statistically self-similar if it can be divided into N disjoint subsets which are like - but not necessarily identical to - the original set. The fractal or similarity dimension of A is then determined from the following relation

$$D = \log(N)/\log(1/r) \text{ or } N = 1/r^D \text{ [3]}$$

where D is the fractal dimension, N is the number of identical subsets into which A has been divided, and r is the ratio of the common size of the subsets to the original set A.

If A is Euclidean n-space, then D is an integer - in fact D = n. If D is not an integer, the set A is said to be a fractal.

The existence of a set with a non integer fractal dimension is illustrated in Figure 1. The figure illustrates the recursive rule for generating the Koch snowflake discussed in [5] and [7] among other places. At each iteration, straight line segments - such as the ones on the left - are replaced by the curve on the right. This procedure is allowed to continue forever. One can calculate that

N = number of replicas of the original set = 4
r = size of the replicas in relation to the original = 1/3

$$\text{so that } D = \log(4)/\log(3) = 1.26....$$

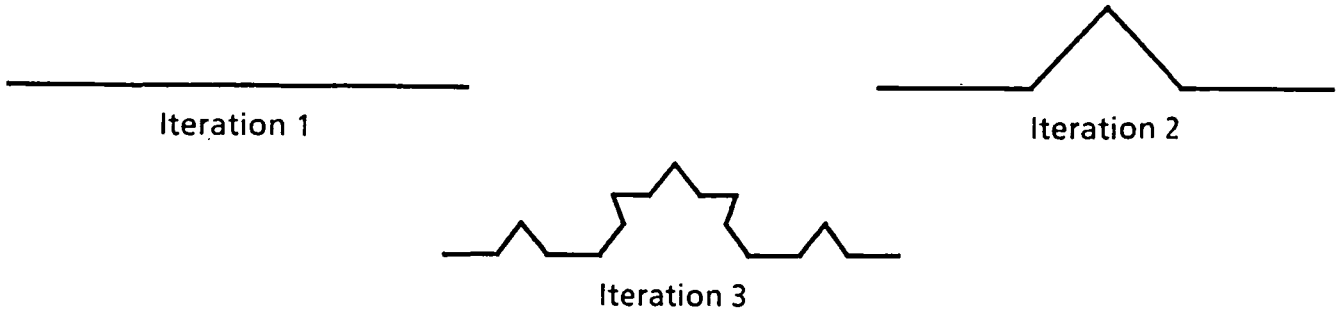


Figure 1

One can also calculate that the length of the curve generated by this process is infinite. Clearly the length of the curve at the kth iteration is 4/3 times its length at the (k-1)st iteration. Thus at iteration k, we have the length, l, given by

$l = l_0(4/3)^k$ where l_0 is the length of the original segment.

The last calculation shows the relationship between the length of a fractal curve and the measuring device used to measure it. If the device is only capable of an accuracy of l_0, we would obtain a length of l_0 when we measured the curve. If, however,the device had an accuracy of $l_0(1/3)^k$, we would obtain a length of $l_0(4/3)^k$. Thus as the measuring device becomes more accurate, the length of the curve increases without bound.

Box Dimension: The example above shows how to calculate the fractal dimension for an artificially generated fractal set. In nature - as well as in some artificially generated fractal sets - the calculation can be far more difficult. This may be due to the complexity of the set or due to the fact that the set has the property of statistical self similarity rather than self similarity.

A useful approximation to the fractal dimension is the box dimension described by Voss in [7]. To obtain the box dimension, the set is covered with a grid of boxes of size r. Let N(r) denote the number of boxes containing a point of the image. Then N(r) is proportional to $1/r^D$. By the appropriate choice of r, Voss shows that one obtains $N(r) = 1/r^D$.

Lacunarity: Mandelbrot showed [3] that the fractal dimension of a set alone was not always sufficient to characterize the appearance of the set. Sets with the same fractal dimension were shown to have different appearances due to texture or graininess. To differentiate between such sets, Mandelbrot introduced the concept of lacunarity which is defined to be

$$L = E((m/E(m)-1)^2) \quad [1, 3]$$

where E denotes expected value and m denotes the mass of the fractal set with some given density function P(m). If the mass is close to the expected value - as in a compact, dense set - this quantity will be small. If the mass is distributed very unevenly this quantity will become larger. Thus lacunarity provides another way of characterizing sets.

Voss [7] provides a way of estimating lacunarity from an image by approximating the mass density function from the image data. Keller [1] provides an alternative yet similar measure of lacunarity. Each of these estimates may be calculated from information contained in the gray scale of the data.

FRACTALS IN IMAGE PROCESSING

In work by Pentland [4], it is demonstrated that a scene may be segmented and classified by computing the fractal dimension of subsets of the image. His approach is based on the Fractal Imaging Theorem which describes the precise relationship between a fractal and the gray scale image of that fractal.

The author used a least squares fit to the Fourier power spectrum over 8 x 8 pixel grids. By choosing appropriate break points in the histogram of the calculated fractal dimensions, Pentland was able to segment natural scenes into their components - e.g., water and land, water land and sky, etc.

In work by Keller et. al. [1, 2], image segmentation was achieved by using the fractal dimension - obtained from the interpolated box dimension developed in the paper - and the lacunarity feature described in the previous section. In these works, regions of like fractal dimension and lacunarity were determined by the k-means clustering algorithm. Segmentation was accomplished by plotting areas with similar characteristics in the same color.

EDGE DETECTION USING FRACTAL GEOMETRY

Using the theory and research described in the previous two sections, work was begun to develop techniques which would permit the segmentation of an image so that manmade objects could be extracted from natural backgrounds. In addition, it was desired that the segmentation be performed more rapidly than is possible with the techniques previously described. The work by Pentland in [4] required the calculation of FFT's, a regression step, and the construction of a histogram of fractal dimensions. Keller's approach in [1] required several passes over the image and the application of a clustering technique before the segmentation was performed.

It was not expected - nor did it turn out - that the overall segmentation achieved by a faster approach would be as good as that achieved in the previous techniques. In fact, the only segmentation desired was a division of the scene into two parts - the part containing the manmade space object and the part containing everything else.

Approach: An existing computer program was used to generate cloud cover which had fractal dimensions of 2.2, 2.4, 2.5, 2.6, 2.7, and 2.8. Images of model toy space shuttles and a model satellite - representing manmade objects - were captured and combined with the clouds to form 256 x 256 pixel images with 16 colors or gray scale steps.

Once the images were created, programs were written to extract the box dimension and lacunarity measurements on grids of size 4 x 4, 8 x 8, and 16 x 16.

Simple threshholding was then performed on the fractal dimensions and lacunarity measurements taken. Thus the program was written to allow the user to input threshhold values and display only those grids whose fractal dimensions, exceeded (fell below) the fractal threshhold and whose lacunarity measurement exceeded (fell below) the lacunarity threshhold. In addition, an m x m grid was displayed (where m = 256/n and n is 4, 8, or 16). A pixel in the m x m grid would be "on" if and only if the corresponding n x n grid passed the fractal and lacunarity threshhold tests. It was this final m x m grid which was sent to a lookup table and ultimately will be sent to a neural network or other pattern recognizer.

Implementation and Improvements: The computer programs described above were written and applied to the data sets. While it was not the case that it was always possible to recover the manmade object from the background completely, it was observed that this approach generally does an excellent job of edge detection on the manmade objects.

It was determined - experimentally - that the 4 x 4 grids worked better than the 8 x 8 or 16 x 16 grids. While the latter two grid sizes performed a good job of edge detecting the manmade objects, they also tended to select boundary areas in the clouds. In particular, it was difficult to find appropriate threshholds which would remove those regions where clouds met regions of clear sky (or open ocean depending on the orientation of the observer to the target).

Using 4 x 4 grids was not without difficulty. The major problem was that when computing the box dimension based on the gray scale of the image, the algorithm described in [7] requires that the box be scaled by the same amount in each of the x, y, and z directions. This allowed for the choice of r to be 1, 1/2, or 1/4. If r = 1 nothing is gained since there is one box. If r = 1/4, N(r) must necessarily be 4 (one box per pixel) and we would obtain D = 1. Thus the only choice was to select r = 1/2. This scale factor, however, allowed for little or no separation in the z or gray scale direction. With 2 vertical boxes, gray scales 0-7 were grouped together and those from 8-15 were grouped together.

To overcome this problem, it was decided to scale by 1/2 in each of the x and y directions and by 1/8 in the z direction. While this gave a distorted fractal dimension when the formula for computing the box dimension was applied, it allowed for better detection of variations in the gray scale. The distorted measurements were converted to more realistic, believable fractal dimensions by using a linear scaling to make them all lie in the interval {2, 3}. This conversion is not necessary from a mathematical point of view, but rather was done for aesthetic reasons.

For the higher fractal dimensions of clouds, it was noticed that occasional stray, isolated pixels were turned on in the 64 x 64 image to be passed into the lookup table. In hopes of achieving a more distortion free image, an erosion step was added to remove stray pixels and a dilation step was inserted to strengthen the detected edges. In the erosion step, any on pixel with no neighbors was turned off. For the dilation step, we used a 2 x 2 square of pixels and performed the dilation as described in [6].

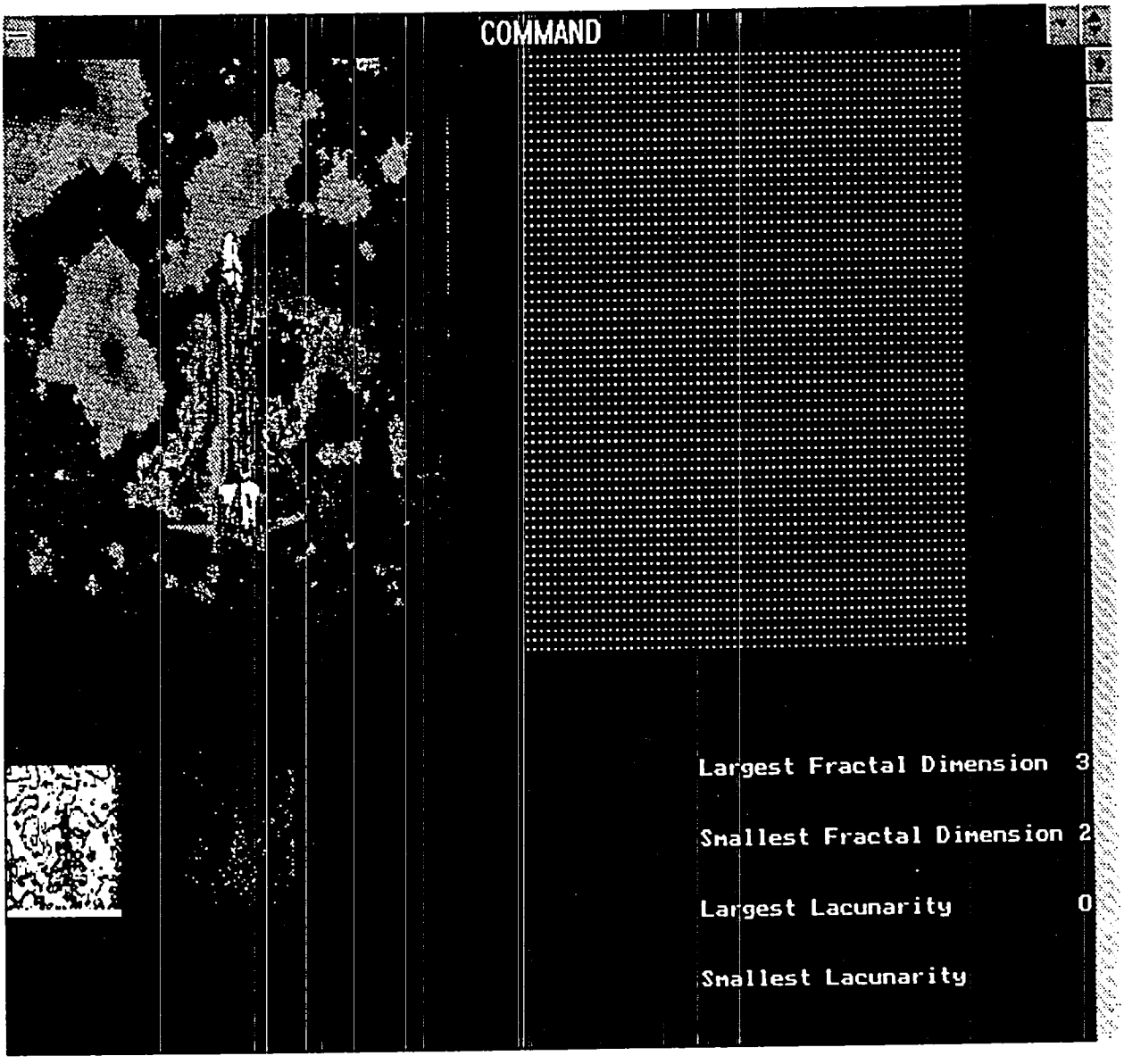

COMMAND
Largest Fractal Dimension 3
Smallest Fractal Dimension 2
Largest Lacunarity 0
Smallest Lacunarity

FIGURE 2

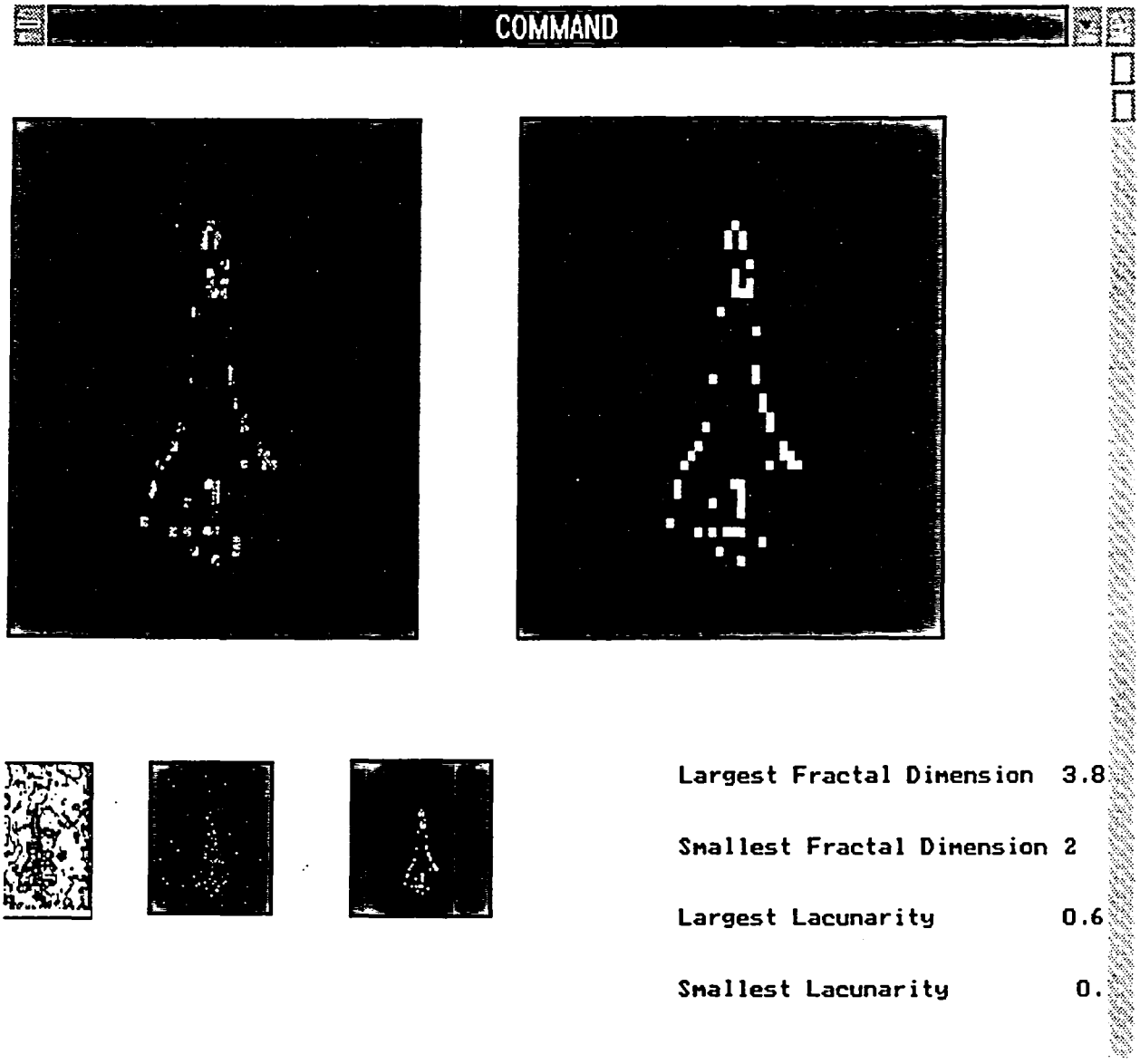

COMMAND
Largest Fractal Dimension 3.8
Smallest Fractal Dimension 2
Largest Lacunarity 0.6
Smallest Lacunarity 0.

FIGURE 3

SUGGESTIONS FOR FURTHER STUDY

Most of the emphasis of this work was placed on the development of the target detection portion of the program. The participants in the work hope to carry the work further and to replace the table lookup described earlier with an appropriate pattern identifier, e.g., neural network. This would allow automatic identification as well as automatic detection of space targets. In addition, the authors believe that other image processing processing techniques - e.g. openings and closings - may enhance the quality of the 64 x 64 image being passed to the pattern recognizer.

Finally, the technique must be verified on real data. While the model targets and artificially generated backgrounds go far to proving the validity of the concept, it should be verified on actual images from space. Furthermore, the proof of concept has been restricted to the open ocean, the open sky, or backgrounds representing a combination of clouds, ocean, and sky. However no provision was been made to test it against a background containing land masses e.g. continents and/or islands.

Cooperation between the authors is expected to continue in the future on the solution of these problems.

REFERENCES

1. Keller J. and Chen S. 1989. "Texture Description through Fractal Geometry" Computer Vision, Graphics, and Image Processing 45. pp 150-166.

2. Keller J. and Chen S. 1990. "Local Fractal Geometric Features for Image Segmentation". International Journal of Imaging Systems and Technology, Vol 2, 267-284. pp 267-284.

3. Mandelbrot, B. 1983. The Fractal Geometry of Nature, W.H. Freeman and Co. pp 25-57, 310-318.

4. Pentland, A. 1988. Natural Computation, "Fractal-Based Descriptions of Surfaces", MIT Press. pp 279-298.

5. Pokornoy, C.K. and Gerald C.F. 1989. Computer Graphics The Principles Behind the Art and Science, Franklin, Beedle & Associates. pp 340-346.

6. Shapiro L, et al. "Computer Vision - A Training Course Presented to NASA at JSC". Seminar Notes from Machine Vision International Training Course 1990. pp 1-45.

7. Voss R.F. 1988. The Science of Fractal Images, Springer-Verlag. pp 21-31, 59-62, 67-69. Editors Peitgen and Saupe.

CLOSED-LOOP HABITATION AIR REVITALIZATION MODEL FOR REGENERATIVE LIFE SUPPORT SYSTEMS

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1991

Johnson Space Center

Prepared By:	Maxwell M. Hart
Academic Rank:	Associate Professor
University:	University of Mary Hardin-Baylor
Department:	Dept. of Mathematics and Physics Belton, Texas 76513

NASA/JSC

Directorate:	Engineering
Division:	Crew and Thermal Systems
Office:	System Engineering Analysis
JSC Colleague:	Chin H. Lin
Date Submitted:	August 9, 1991
Contract Number:	NGT-44-011-800

ABSTRACT

The primary function of any life support system is to keep the crew alive. This means it must provide for breathable air, potable water, edible food and disposable waste. The interesting thing is that none of the chemical elements are ever used up. They just change from one molecular form to another. In a well-balanced or "regenerative" life support system, the various components are each using what is available and producing what is needed by other components so that there will always be enough chemicals in the form in which they are needed. There would be very little excess and there would never be a long-term build up of excess compounds.

Humans are not just users of the system, they are actually one of the participating components of the system. Since humans are involved, it is desirable for the system to be self-sustaining and independent of outside resupply. If this could be achieved then the original chemicals would be continually recycled and the humans would not be jeopardized by resupply schedules. This would make the system virtually a Closed-loop Habitation (CH). Our earth is one of the best examples of a CH. When we send astronauts on extended space voyages such as to the moon or Mars we would like to provide them with a scaled down miniature CH.

There are a number of inherent difficulties in trying to create a miniature version of our earthly CH. 1) Earth's size gives it sufficient mass and sufficient gravity to retain its atmosphere. A scaled down version must use artificial containment such as sealed enclosures. Thus it is subject to leakage. 2) The earth has a very large, very diverse and complex ecosystem that would be difficult to duplicate on a small scale. 3) The size of the earth's atmospheric buffers would be prohibitively massive in a proportionately scaled down version for even one person, much less an entire crew, on a space voyage. 4) Natural disasters such as violent weather or wildfire are necessary to the earth's ecosystem but would be hard to duplicate and difficult to manage on an isolated lunar or Martian outpost. The list goes on and on.

In a miniature CH we must provide a minimal structure and make up the difference by artificial components such as physicochemical systems that perform the conversions that the earth can achieve "naturally". To investigate the interactions of these components we have designed a computer model that simulates a miniature CH with emphasis on the air revitalization components. It is called the Closed-loop Habitation Air Revitalization Model (CHARM).

INTRODUCTION

Present life support systems that rely heavily on expendable materials are becoming unattractive for extended space voyages (e. g. Moon or Mars). Strict mass and volume constraints make it necessary to consider regenerative life support systems as an alternative. Present technologies, including physicochemical and biological processes, have demonstrated the feasibility and potentials of regenerative life support at a component level. However, there is much work that remains to be done to advance the maturity of the technologies and to study integration of the regenerative life support components. NASA has identified regenerative life support technologies as one of the critical areas of research and development for space exploration. Not only does the success of the mission depend on life support, so does crew existence.

METHODOLOGY

This investigation involves the development of a computer model that simulates the operation of the air components of a regenerative life support system. There are several modes in which the model can be run to aid the study of regenerative life support systems.

Verification Mode

The model is run behind an actual life support system test bed that it models. In this mode the model can be used for any one or more of at least three different purposes:
1: to validate model correlation with test bed;
2: to corroborate test bed results; and
3: to understand and explain unexpected test bed results.

Warning Mode

The model runs ahead of an actual life support system test bed that it models. In this mode the model can be used to predict what the test bed will do, thus eliminating any surprises that might occur. The researcher can then be ready with the extra equipment, supplies and procedures that might be necessary for the upcoming event.

Preparation Mode

The model is run for many weeks testing the boundary conditions in preparation for the construction of a new test bed facility. This could save a great deal of expense in construction costs by making sure the new facility can handle all of the tests for which it is being designed.

Investigation Mode

The model is run independently of any test bed facility. In this mode the model can run test for which there is no test bed available. Unexpected results could pinpoint areas for future investigation.

Blanket Mode

The model runs large amounts of data to cover the gaps between the data values obtained from a test bed.

Query Mode

The model can be used to search for particular nodes of confluence or the occurrence of particular events.

PARAMETERS

One way to explain the operation of CHARM is to list all of the parameters and explain which ones can be varied and which ones must remain fixed and at what value.

System Constants

These values are well known constants.

A. Atomic weights
   (Based on IUPAC Atomic Weights of the Elements 1981)
   1. Hydrogen = 1.00794
   2. Carbon = 12.01100
   3. Nitrogen = 14.00670
   4. Oxygen = 15.99940

B. Conversion factors
   (CRC Handbook 1987)
   1. Lbs. to kg = 0.45359237
   2. Ft. to m = 0.3048

Crops

A. Any area ratio of the following four crops can be grown: lettuce, potato, soybean, wheat. The current area is 9.75 m^2.
B. Growth characteristics are totally adjustable to obtain desired curve.
C. Any amount of fallow area can be designated.
D. Each crop can be set independently for any ratio of daylight vs night. This affects growing time, but growth curve must be set separately. Currently, potatoes have 24 hrs. daylight, others have 16 hrs.
E. Each crop can be set independently for harvest time. Currently, lettuce is harvested in 30 days, potatoes in 160 days, soybean in 90 days, and wheat in 80 days.
F. All crops are checked to maintain mass balance.

Growth Chamber Parameters

A. Volume is a variable affecting leakage. Current value is 25 m^3.
B. Growing area is a variable. Current value is 9.75 m^2.
C. Air pressure is a variable. Current value is 10.2 psia.
D. Temperature is a variable affecting density. Current value is 20 degrees C.
E. Light intensity is constant for each crop type:
   Lettuce 450 micro mol/m^2-s
   Potato 300 micro mol/m^2-s
   Soybean 700 micro mol/m^2-s
   Wheat 1200 micro mol/m^2-s.
F. When potatoes, soybeans and wheat are harvested, the inedible portion is discarded. When lettuce is cut, the inedible portion continues to grow new crop.

Crew Cabin

A. Volume affects leakage. Current value is 30 m^3.
B. Crew number is variable. Current value is 1 person.
C. Average crew weight is assumed to be 155 lbs. for diet.
D. Energy usage is 2700 kcal/man/day for diet calculation.
E. Crew diet is a variable. Current values are from the NIH recommended diet.
F. Stored food is automatically used as needed.
G. Uneaten edible portions of grown food is automatically computed, based on chosen diet, removed from system, and added to waste.
H. All output is checked for mass balance.

Buffers

A. There are ten fluctuating variables that are tracked. They are: nitrogen, oxygen, carbon dioxide, water, hydrogen, methane, hydrazine, food, waste, and resupply.
B. The volume of the gas storage buffers can be adjusted anywhere from zero to any positive real number.
C. There are four sensor variables on each buffer. They are: empty, min., max., and full. Each can be adjusted individually as needed. The model is programmed to run physicochemical systems when values go out of the range between max. and min. The model will use resupply and auxiliary storage when values try to go below empty or above full.
D. Besides the effect of the various components on the buffer levels, there is also provision for external factors that might affect the buffers. These can be used in addition to, in place of, or instead of the human component, the plants, or both.
E. The variable leakage rate can be adjusted anywhere from 0 to 100 percent. All or a portion of the leakage can be made to reflect any of a variety of factors. These include the relative abundance of each atmospheric gas in the growth chamber and in the crew quarters separately along with the volume of each. It can also reflect the relative size of each molecule in case of permeable membrane type leakage, or each molecule leakage rate can be set individually.
F. There are four physicochemical systems used by the system. They are:

Sabatier	$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$
O_2 generator	$2H_2O \rightarrow O_2 + 2H_2$
N_2 supply	$N_2H_4 \rightarrow N_2 + 2H_2$
CH_4 burner	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$

These are controlled by the user by changing the buffer sizes mentioned above in B and C.
G. The length of the calculation cycle can be varied to allow for more precise checking. The system currently operates on an eight-hour cycle.
H. The program time span depends on the length of the calculation cycle in G above. On an 8-hour cycle, the program runs for one month. However, several consecutive months can be run at a time with perfect compatibility.
I. There is an end-to-end mass balance check that can be performed to insure against unexplained loss or gain of mass throughout the program span.

DISCUSSION

This program can produce data in a variety of graph and table formats. For most purposes, the preferred form is graphical. A few graphs have been included in this report to help illustrate the scope of CHARM. Figures 1 through 4 show the growth curves for four crops generated by CHARM. For this report we are using, with only minor adjustments, the formulas and data of Volk and Cullingford, "Crop Growth and Associated Life Support for a Lunar Farm," 1989.

Figure 5 is a graph showing the storage buffers of a system similar to the facility now producing lettuce in the "ten-foot chamber" at JSC. For this simulation, CHARM was provided with very large buffers. This prevents the built in physicochemical systems from going into effect, trying to keep the system in range. To produce this graph, the system grew lettuce on a 16 hour/day cycle of daylight vs. 8 hours of night for 30 days and then harvested.

The data in this graph, as in many of the others, reflect the amount of fluctuation in each compound. There is no indication as to how much of this variation is affecting the atmosphere or how much is being added to, or removed from the atmosphere as compensation for these internal changes. At the present time, because this information is not yet available, it is assumed that the atmosphere is remaining of constant composition and all variation occurs only in the storage buffers.

The system was arbitrarily assigned a leakage rate of 10% for carbon dioxide 5% for oxygen, and 5% for water vapor. The system was given a 0% leakage rate for nitrogen because the 10-foot chamber, being at ambient pressure on Earth instead of the airlessness of space, has about the same amount of nitrogen leaking in as it does leaking out. This situation will change somewhat when the chamber is lowered to 10.2 psia. Then, more air will be leaking in then out. CHARM can also be run at negative leakage to simulate this result. For the present, the other gases, besides nitrogen, will have a tendency to leak out, however, since they are at higher concentration in the chamber than similar gases outside the chamber. Figure 6 is a graph similar to figure 5 but it represents a facility growing multiple crops. The fraction that each crop represents of the total growing area is given above the figure.

Several facts are apparent from figures 5 and 6. It can be seen that more oxygen is being produced by the plants than is being consumed by the human in the model. Secondly, the plants are using more carbon dioxide than is being produced by the human. Thirdly, and most surprisingly, the water level is going down. This only represents the chemically available water. Atoms from the water are being incorporated into living tissue in the plants faster than water is being produced by oxidation of food by the human.

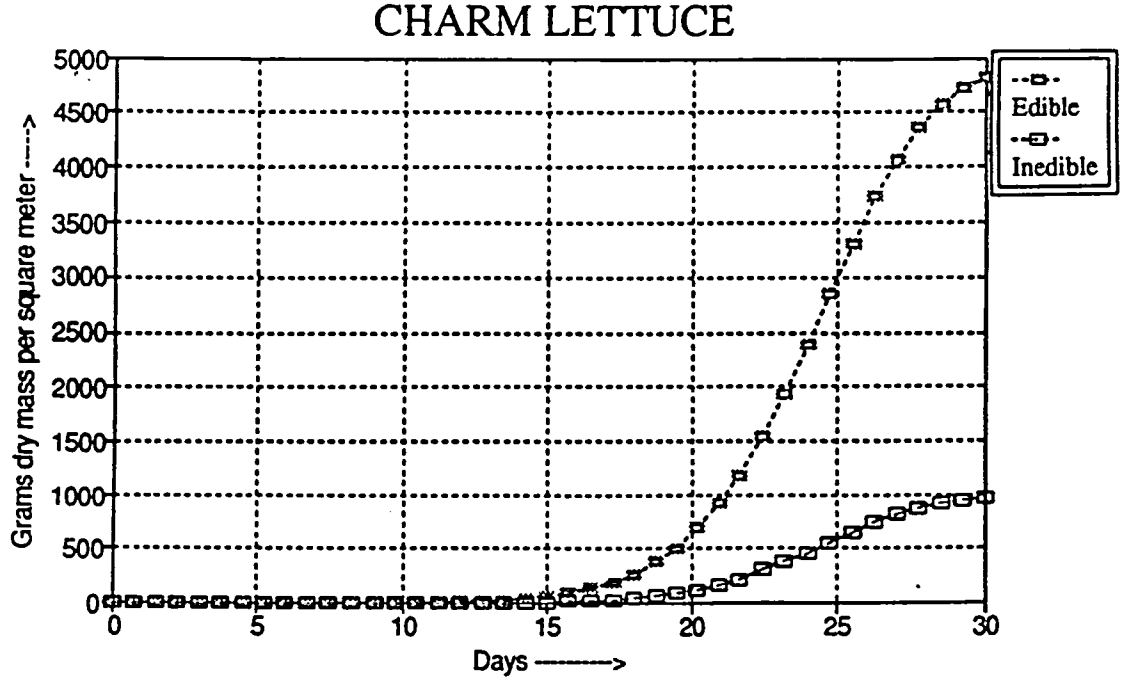


Figure 1.-Thirty day lettuce growth curves.

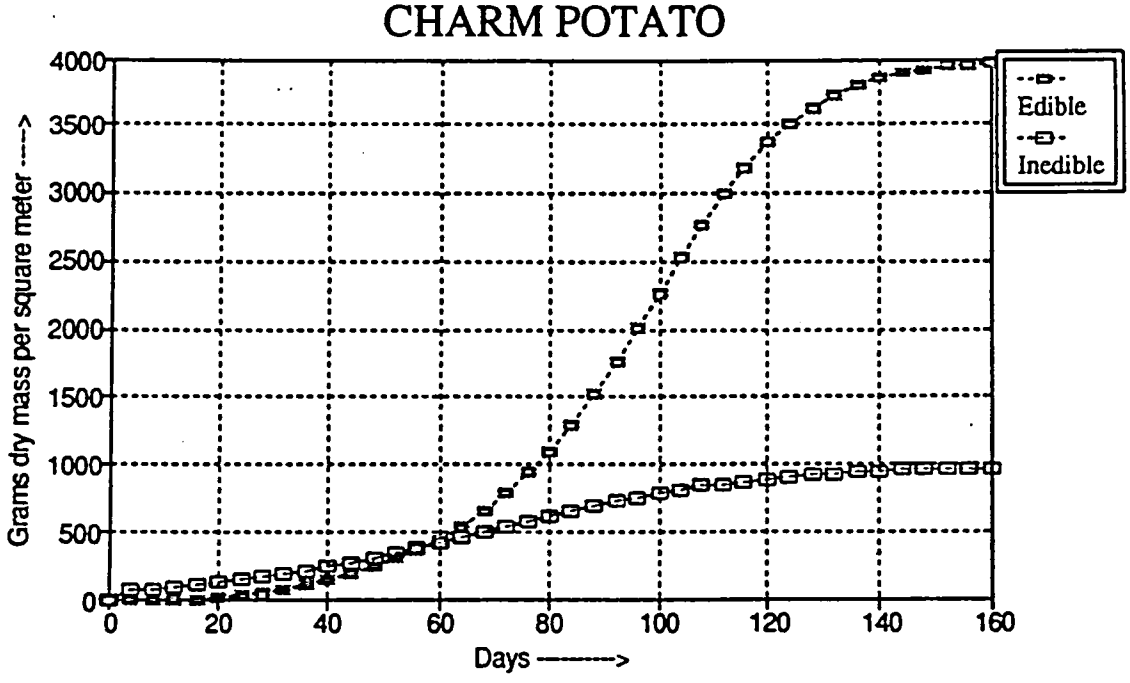


Figure 2.-One hundred sixty days of potato growth.

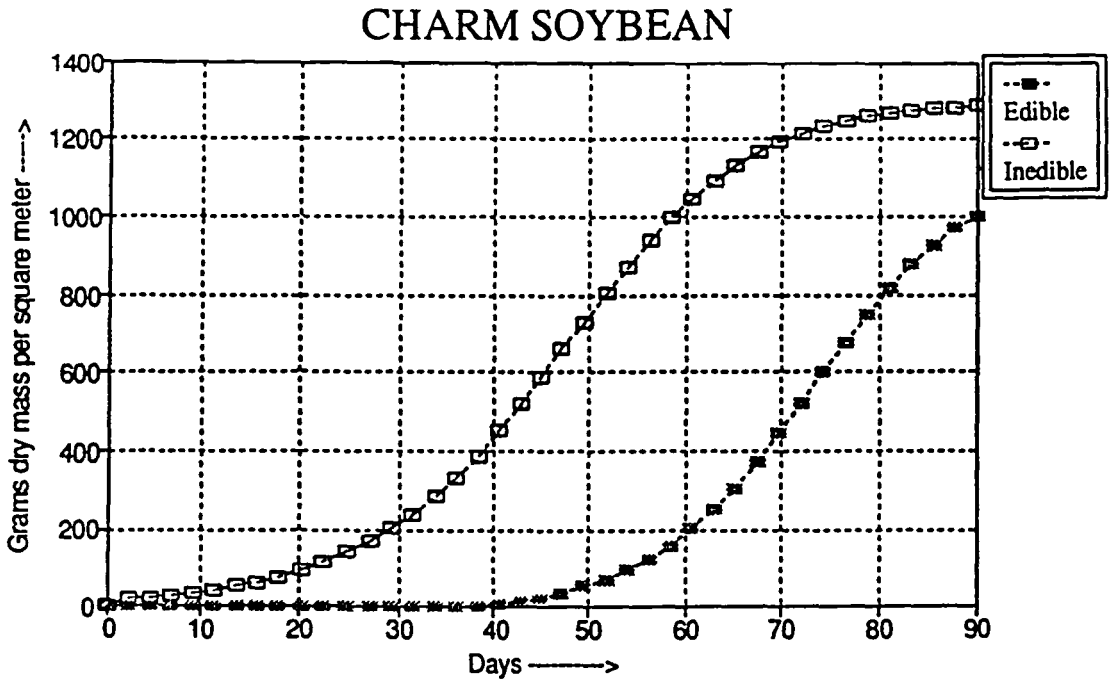


Figure 3.-Ninty day soybean growth curves.

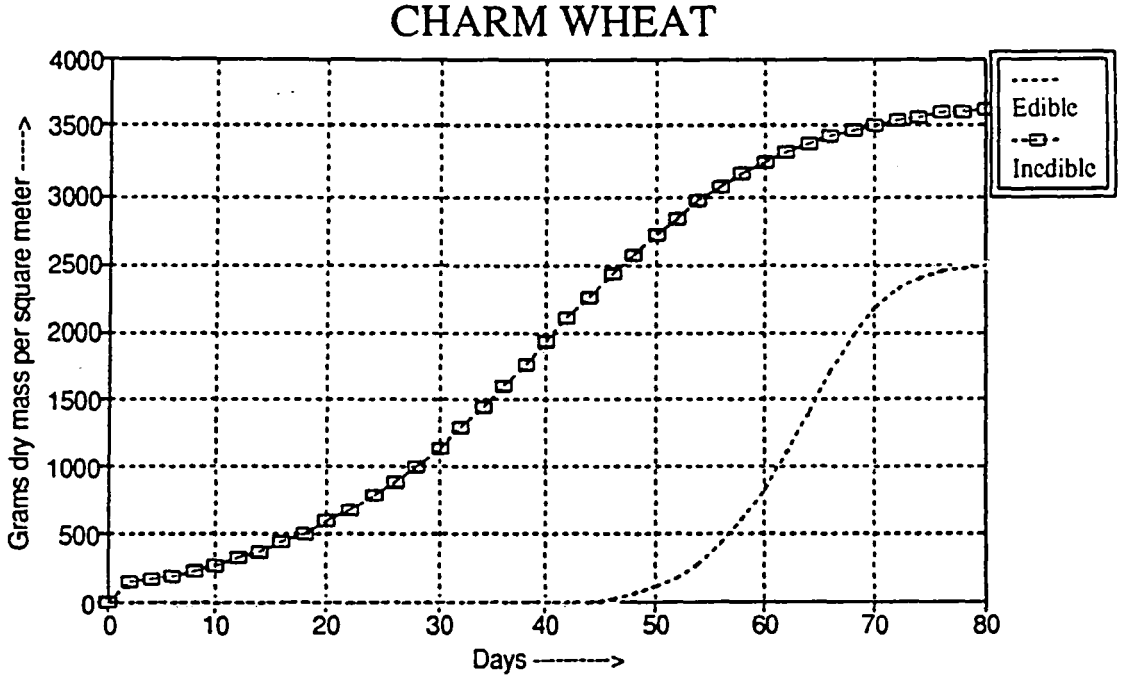


Figure 4.-Eighty day wheat characteristic curves.

Crop Ratio: 1.0 Lettuce, 0.0 Potato, 0.0 Soybean, 0.0 Wheat
Leak Rates: Nitrogen 1%, Carbon Dioxide 1%, Oxygen 1%, Water Vapor 1%

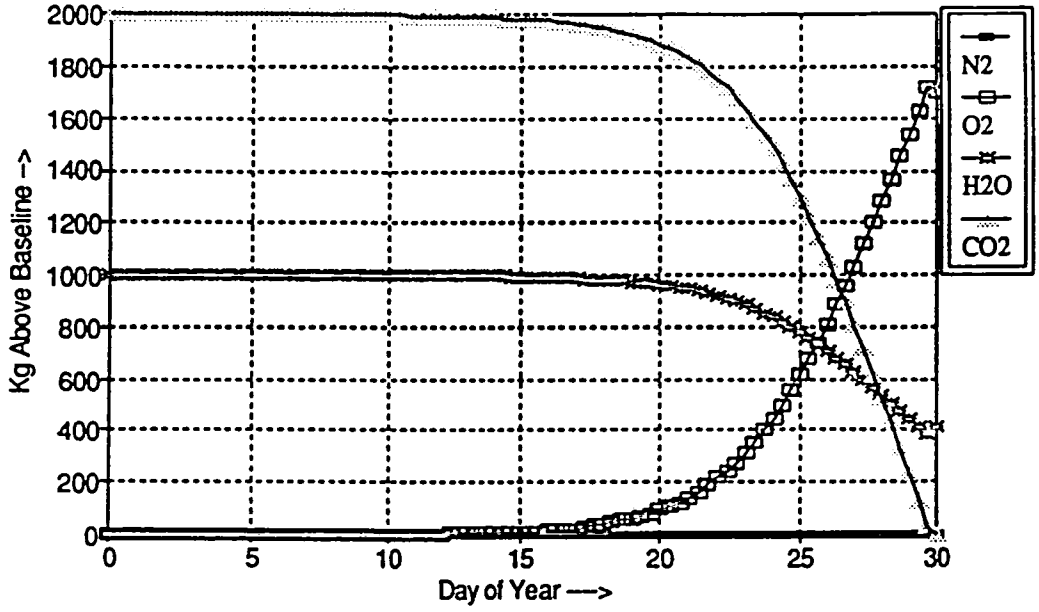


Figure 5.-Flux for 31 days of lettuce growth.

Crop Ratio: 0.0 Lettuce, 0.2 Potato, 0.7 Soybean, 0.1 Wheat
Leak Rates: Nitrogen 1%, Carbon Dioxide 1%, Oxygen 1%, Water Vapor 1%

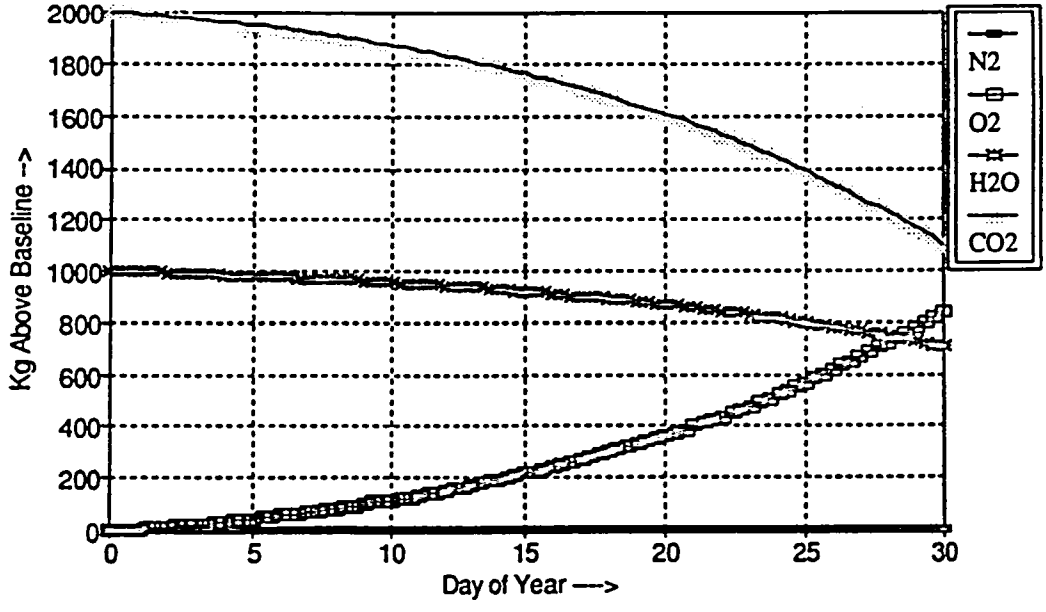


Figure 6.-Flux for 31 days of multiple crops.

Plants release a large amount of transpiration water into the atmosphere. This is a physical phenomenon which, although important, is not related to the chemical depletion of water by the plants. The missing component in this cycle is waste disposal. When left-over food, inedible plant mass and other waste is processed, some of the excess oxygen is consumed. Waste management also helps to make up some carbon dioxide and water losses. There is an inherent problem in the modeling of only one component of a complex system. The lack of a complete cycle makes verification difficult and somewhat questionable. One of the very nice features of CHARM is that it can easily be expanded to incorporate more components until it eventually models a complete regenerative life support system.

Figure 7 demonstrates the difficulty of the leakage question. In this example, 10% of the carbon dioxide is being lost per day compared to the rather low leakage rate of 1 percent for all of the other examples. Ten percent sounds very high because, at this rate, the carbon dioxide should be completely replaced every ten days. This is misleading. Compare the graph of figure 7 with that of figure 5. It is difficult to tell the difference. Actual leakage is 10% of only the atmospheric carbon dioxide. More carbon dioxide is in buffer storage. Storage buffers are assumed to be well sealed with hardly any leakage. Although the nominal rate may be 10% of atmospheric gasses or even higher, the effective rate on the system depends heavily upon buffer sizes. If storage buffers are a relatively small component of the total atmosphere, effects of leakage can be much greater. It sounds trite, but it is important to note: the more replacement mass you have in storage, the less you notice leakage. This is one of the factors that must be considered when deciding how large these buffers should be. If leakage in an actual test bed is high, this could make verifications difficult. Although it is easy to calculate how much has leaked, it is difficult to tell exactly when, where, and in what form it leaked. Many claims can be made with the assertion that any discrepancy is simply due to leakage. A computer model that simulates leakage can help to verify or refute some contentions.

Figure 8 models a life support system similar to that for figure 5 except that the buffer sizes are much smaller. Automatic physicochemical systems try to keep the system operational between relative maximum and minimum values. Absolute boundaries of full and empty are maintained by processing and storage of excess supplies in other forms such as fuels and liquid water. Shortages are made up by use of resupply stockpiles. All processing and resupply is tracked and available for study. The interior, relative, boundaries are maintained by physicochemical systems as long as the necessary chemicals are available.

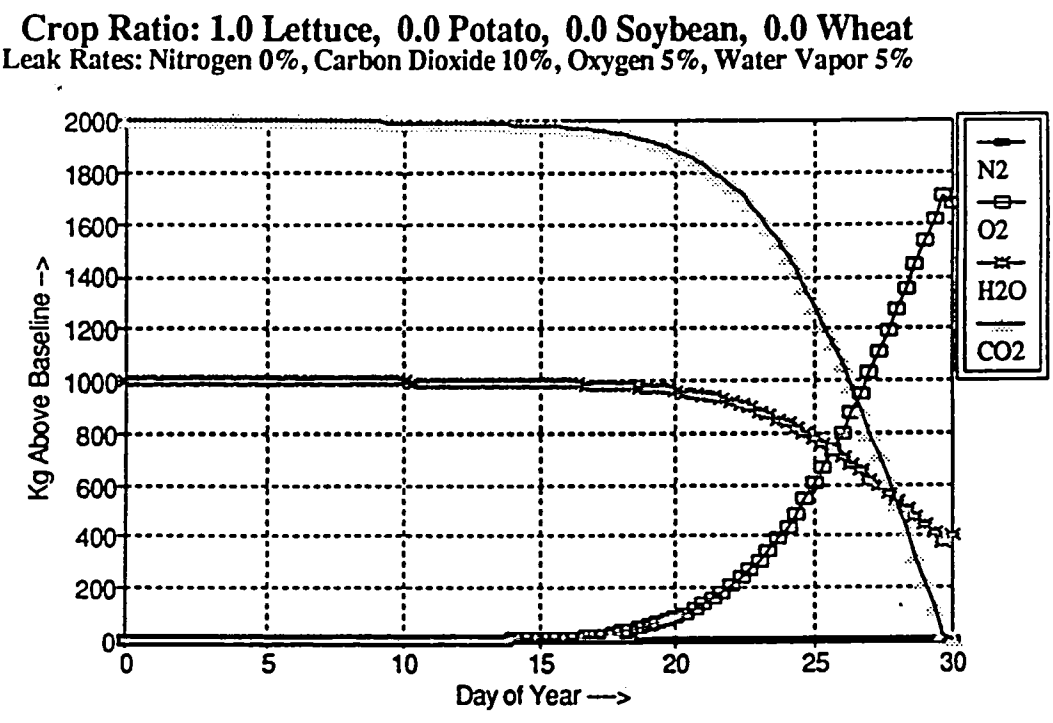


Figure 7.-Moderate leakage rate.

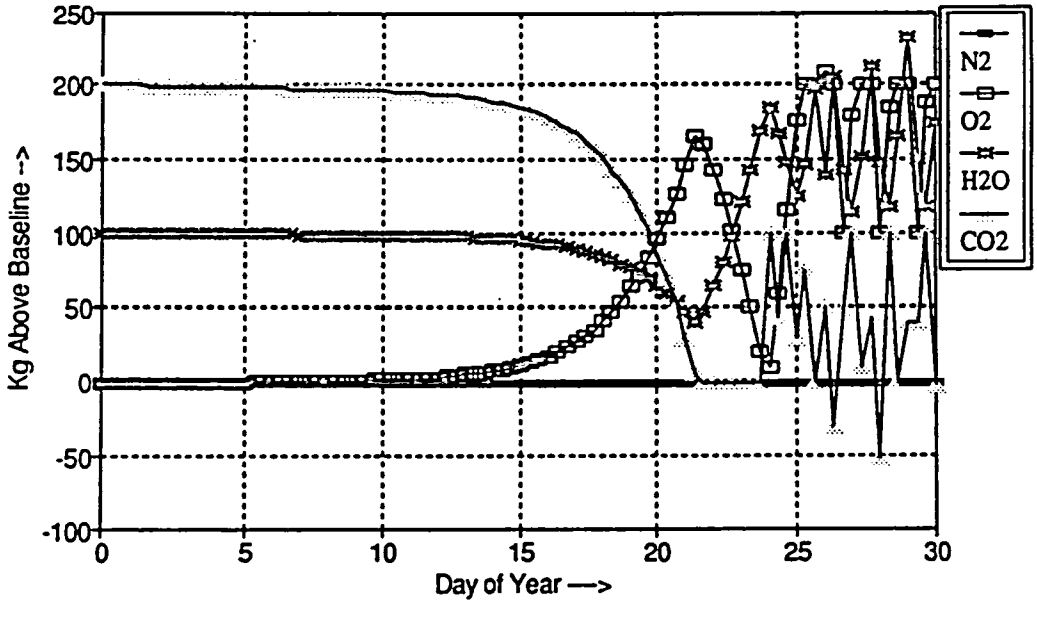


Figure 8.-Small buffer volumes.

The final month of this summer project was spent in the development of a user interface for CHARM. The purpose of this interface is to make CHARM user friendly. It is desirable that this program be useable by someone who is not familiar with Quattro Pro and even by those who do not normally use a spreadsheet in their work. The framework for this interface was built into the program during initial development so that it was a simple matter of tying it all together into one coherent package. To accomplish this, three new programs were written. These will be described in the three following sections.

Heading

This program introduces users to CHARM. It allows them to log on and determines if they are new users. This information is used in two ways. It is incorporated into the printed records along with the current date and time as a record for user reference. It is also used by the program to allow regular users to skip some of the constant messages and menus that only serve to slow down the program for those who no longer need them.

This program also contains the main menu. From this menu the user is able to choose the type of work to be done. The user can scroll through the menu by any of three methods: the space bar (down only), the arrow keys, or the mouse. As the user highlights a selection, a short description appears at the bottom of the screen. Once the decision has been made, the user has three methods of making the selection: highlight the selection and press the enter key, use the mouse and click on any entry even if not in highlight, or type the first letter of the selection (the first letter is in red, the rest are black).

When users finish particular tasks, program execution returns them to the main menu. The bottom choice on the menu allows users to exit the program and automatically logs them off. This feature of returning to the most recent menu is common in most of the other menus throughout the user interface except where there is no further need for some menu. In that case, the user is returned to the previous menu.

Another decision on the main menu allows the user to select which printer to use. The remaining choices, besides the ever present "Help" selection which allows the user to call up a help screen, have to do with deciding whether to do input, processing, or output. Each of these latter three selections also require the user to make other decisions about the type of input, processing, or output that is to be done. Much of this additional selecting is also handled by the "heading" program in a manner similar to that described above.

Files

This is the second of the three human interface programs of CHARM. Its primary function is to shuttle the various files in and out of main memory. It determines exactly which files will be needed, opens the files and leaves them open just as long as necessary. If a file will no longer be needed, it is closed and removed from main memory to make room for other files. This function is very important because, if too many files are open at one time it will cause the program to terminate for lack of memory space.

With Quattro Pro it is not necessary to open a file when information must be retrieved from it. Every time a file is opened, it can be updated with all pertinent information from all other files, even if they themselves are not open. However, there are many procedures that require a particular file to be open. Of even greater importance, a file cannot be updated if it is closed. Therefore it is necessary to open a file and update it before it is used so that other files will not receive old data from a closed file that was not updated in time. It is the function of this "files" program to hand the user over to the "update" program when a file may contain old data. A special benefit of this opening and closing of files is that files are always being "saved" so that abnormal termination of the program or system failure will never result in much lost data.

Update

It is the function of this program to update the files in the proper order so that no file ever receives old data. This program works in cooperation with the "files" program to guarantee data fidelity. This program also works with "files" in the area of data retrieval. As files are being updated, they become available for output such as in the graphing routines. This results in a considerable savings in time. Since the files must be loaded and unloaded anyway, it makes sense for the computer and printer to work in tandem so that the printer can be working at the same time that the computer is processing. Although more graphs are available and many more are possible, the program currently allows up to eleven different graphs to be printed for each month of data for the entire run of the experiment. The program even prints out a data sheet showing the current values of the user chosen parameters, the date and time of the run along with the name of the user. Once the user has chosen the variables, dates of the run dates of crop failure and graphs to be printed, the program will run for as long as it takes to print all of the graphs on the chosen printer. No additional user intervention is necessary except possibly supplying more paper to the laser printer.

CONCLUSION

CHARM is incorporated within a spreadsheet, Quattro Pro. This aspect gives it some disadvantages. Because of its dependence on Quattro Pro, CHARM can only be used in the Quattro Pro environment. This limits its accessibility. CHARM also uses a very large amount of main memory. When used on a machine with limited memory, Quattro Pro must constantly move files in and out of disk storage. This slows down processing considerably. If CHARM were used on a system where it had access to 1M or 2M of memory it would be able to operate several magnitudes faster. Some operations which are now quite complex would become simple by comparison. This would allow even greater tasks to be performed which are now only barely imaginable. CHARM is also starving for data. There is so much more it could accomplish if it were fed information concerning the effects of varying other values not now incorporated into CHARM.

However, the result of developing CHARM from Quattro Pro is not at all unilateral. There are also some clear advantages. The most obvious advantage is its facility for relatively quick development and rapid changes in its structure right down to its very foundation. Because Quattro Pro is so well tested, readily available, and popularly used, it is virtually error free. This greatly increases reliability and versatility of any imbedded programs. CHARM runs very fast for this type and size of program and allows for many variable parameters. Parameters which are not, now a part of it can be added if the relationships are known. Some unknown relationships can be investigated by varying parameters to extrapolate from the overlapping relationships which are incorporated within CHARM. The prototype nature of this program within a spreadsheet belies its power and functional status. The entire program, actually nineteen separate spreadsheets and programs plus several other utility programs, was written in just six months of programming by a single programmer. This fact is a testament to its potential. It can only be imagined what could be incorporated into CHARM with sufficient input and continued mathematical enhancement. One final observation. It is probably fortunate that CHARM was not developed on a larger computer. If it had been, it would have incorporated all of the power of that machine and become a much more powerful program. But as a trade off, it would also have become dependent on that larger machine and therefore would no longer be compatible with any of the much more numerous computers which are widely available around the country. It is impossible to predict how many people will ever see or use this program, but as much effort as has gone into it, I do hope someone, somewhere will find a use for it. In any case, it has been a good experience and a great project. For more information please call Chin H. Lin 483-9126 or Donald L.Henninger 483-5034.

PHOTOGRAPHIC IMAGE RESTORATION

Final Report

NASA/ASEE Summer Faculty Fellowship Program-1991

Johnson Space Center

Prepared by:	Gerald E. Hite
Academic Rank:	Associate Professor
University & Department:	Texas A&M Univ. at Galveston Dept. of Marine Science Galveston, Texas 77553-1675
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Solar System Exploration
Branch:	Space Shuttle Earth Observation Office
JSC Colleague:	David E. Pitts
Date Submitted:	August 10, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

Deblurring capabilities would significantly improve the Flight Science Support Office's ability to monitor the effects of lift-off on the shuttle and landing on the orbiter. This summer a deblurring program was written and implemented to extract information from blurred images containing a straight line or edge and to use that information to deblur the image. The program was successfully applied to an image blurred by improper focussing and two blurred by different amounts of motion blurring. In all cases the reconstructed modulation transfer function not only had the same zero contours as the Fourier transform of the blurred image but the associated point spread function also had structure not easily described by simple parameterizations. The difficulties posed by the presence of noise in the blurred image necessitated special consideration. An amplitude modification technique was developed for the zero contours of the modulation transfer function at low to moderate frequencies and a smooth filter was used to suppress high frequency noise.

I. Introduction

Photographs of lift-off and landing of the shuttle are taken by the Flight Science Support Office. These are used to provide immediate identification of any potential damage to the orbiter as well as to identify sources and effects of any debris created during these crucial periods. The Space Shuttle Earth Observation Office is the repository of photographs taken by astronauts of environmental, geological, meteorological and oceanographic phenomena observed in orbit. Inadvertently many photographs are blurred by improper focusing, relative camera-object motion etc. To gain access to important information contained in such photographs, it is necessary to apply image processing techniques. The more sophisticated of these involve digitization of the raw images and subsequent computer processing of the digitized images or their Fourier transforms.

The research this summer involved developing and implementing a program that would extract information on the point spread function (psf) directly from any blurred image containing a straight line or edge. Although the technique would work for more general blurrings it was especially useful for the two most common types of blurring: improper focusing and motion blurring. In the case of motion blurring the intensity at a point is spread along a line. Although it need not be uniform along the line, for simplicity it was assumed to be along a straight line. In the case of improper focusing the intensity of a point is spread to neighboring points with the relative intensity depending only on the distance from the initial point. The distribution of intensities of each point in the "sharp" image to the points in the blurred image is given by a convolution sum of the "sharp"image and the point spread function. The advantage of a Fourier Transformation (FT) is that it reduces this sum to a product. If there were no noise, the FT of the desired image could be obtained by simply dividing that of the blurred image by that of the psf. However, for the cases of interest the Fourier transform of the psf (MTF) has contours of zero amplitude and becomes small at high frequency. Consequently noise in frequency regions where the MTF is small would result in serious degradation of the signal to noise ratio. Two techniques were applied to minimize such degradation. The first is the implementation of a two parameter high frequency filter. The second is the imposition of a maximum to minimum ratio limit for the amplitude of the filter. This is important at low to moderate frequencies. There are many excellent tests dealing with digital processing (1,2,3) and discrete Fourier transforms (4,5) that provide background for the work described in this report.

II General Approach and Theory

II.1 Convolution of the Point Spread Function and Noise

In general a blurred image is the result of light that would have arrived at a particular point being spread over a number of points. If the distribution is the same for

all points in the image and if the light is incoherent then in the absence of noise the final intensity at a point (pixel) located by $\vec{m} = [m,n]$ would be given by

(1a) $$f^b_{\vec{m}} = \sum_{\vec{n}} p_{\vec{m}-\vec{n}} \, f^s_{\vec{n}}$$

where p is the point spread function, i.e. the image of an ideal point source in the object plane, $f^s_{\vec{n}}$ is the sharp image that would have formed in the absence of blurring and noise.

For the images of interest the digitization process yields an image made up of NxN pixels (N=512). due to the blurring bringing light rays into the image region that they would have otherwise not have entered it, the summation in this expression should include m's outside the range of 0 to N-1. Nevertheless, the sum is usually assumed to be limited to this range to avoid having to deal with an"underdetermined problem"(6).

Noise is always present in the image. It can be the result of electronics, inhomogeneous chemical processes in the original film, etc. Noise can be divided into two contributions: one independent and the other dependent on the signal, i.e. f^s(7). In the latter case it could be incorporated into the point spread function. Since we have little knowledge of the nature of the noise, except that it is usually dominant at high frequencies, we will describe it by an unknown function $n_{\vec{m}}$ and write the expression for the blurred image as

(1b) $$f^b_{\vec{m}} = \sum_{\vec{n}} p_{\vec{m}-\vec{n}} \, f^s_{\vec{n}} + n_{\vec{m}}$$

The next step is to invert the sum in equation 1 and solve for f^s. Since the summation is usually referred to as convolution, the inverse process is referred to as deconvolution. The basic technique is to expand all functions in terms of a complete set of basis functions which will turn the summation into a simple product of the expansion coefficients.

II.2 Discrete Fourier Transormation

For simplicity we use the traditional discrete Fourier expansion

(2a) $$f_{\vec{n}} = \sum_{\vec{k}} F_{\vec{k}} \, z^{\vec{k}\vec{n}}$$

where $z = \exp(i2\pi/N)$.The inverse Fourier transformation is then given by

$$F_{\vec{k}} = \frac{1}{N^2} \sum_{\vec{m}} f_{\vec{m}} \, z^{\vec{k}\vec{m}}$$

In the following we will use lower case letters, e.g. f, for the spatial images and upper case letters, e.g. F, for the Fourier transformed image. The Fourier transform of the point function, psf, is called the modulation transfer function, MTF.
In terms of the Fourier tranformed function Eq. 1 takes the form:

(3a) $$F_f^b = P_f F_f^s + N_f$$

Deleting the common subscripts and solving this equation for the Fourier transform of the sharp image function yields

(3b) $$F^s = F^b/P - N/P$$

If the modulation transfer function and the noise, N, were both known then equation 3b could be used to find F^s and a subsequent inverse Fourier transformation would yield a sharp image. For the sake of discussion let us assume that we have a reasonably good understanding of the blurring process and have a good approximation to psf and thus to the M.T.F. Most likely the blurring extends over a few spatial pixels and MTF will be small at high frequencies, i.e. large values of k. Blurring due to improper focusing, relative motion, atmospheric turbulence are in general of this type. The first two sources of blurring have MTF's that vanish on contours in k-space. Where ever the MTF is small the contribution from noise, N/P, in equation 3b can be significant. Consequently, approximating F^s by F^b/P can result in large noise contributions. The normal procedure to avoid this problem is to multiply F^b by a "filter" function which uses information about the noise to approximate 1/P where noise is unimportant and avoids contributions from regions where noise dominates. The Wiener filter is the most common filter discussed in the literature (2,3,4,6):

(4) $$W^{-1} = P(1 + |N/F^sP|^2)$$

This expression requires knowledge of the amplitude of the function to be calculated and the unknown noise contribution and cannot be used without further approximation. A simple method to avoid over emphasis of noise is present in section IV.

II.3 Point-Spread-Function for Improper Focusing and Motion Blurring

The point spread function describes how the intensity of any point of the "sharp" image is spread among neighboring points in the blurred image. In the case of improper focusing, the point spread function is angle independent and depends only on the distance from the initial point to neighboring points,

In the simplest parameterization p is constant up to a given radius and zero for larger radii. An annulus of a given thickness could be added to allow the intensity to decrease

(5a) $$0 \le P^{(R)}_{\vec{m}} = P(|\vec{m}|)$$

smoothly to zero.

In the case of linear blurring in the direction $\vec{m}_B = [m_B, n_B]$, one must follow the path of a square pixel containing the original point in the sharp image. Since the pixel has finite dimensions pixels whose centers lie near but not necessarily on the line of blurring will receive contributions. In general these points will be located by a vector $\vec{m}_q$ which is the sum of a multiple of $\vec{m}_B$ and a vector of at most unit length perpendicular to $\vec{m}_B$. Thus

(5b) $$P^{(L)}_{\vec{m}} = \sum_q l_q \delta_{\vec{a},\vec{m}_q}$$

where q in the sum labels each such pixel center, and l_q (≥ 0) is its intensity, and $\delta_{a,b}$ is the Dirac delta function.

III Construction of the MTF from Blurred Images

III.1 Use of Zero Contours to Parameterize MTF's

In the most common types of blurring, i.e. due to improper focusing and/or motion blurring, the Fourier transform of the point spread function, the so-called modulation-transfer-function MTF, has contours in the two dimensional frequency space along which it vanishes. In the cases where the blurring is uniform, i.e. the intensity in each square pixel is constant or zero, the MTF's for the cases of interest is the Airy function, $J_1(\alpha f)/\alpha f$ or the sinc-function $\sin(\alpha f)/\alpha f$, i.e. the diffraction patterns for a circular aperture and a slit. Clearly these functions oscillate in sign and have their first zeroes at 1.22π and π respectively. Usually the FT of the blurred image, F^2,allows visual identification of the nearest and perhaps other zero contours resulting from the blurring. In the case of motion blurring those parallel contours are more useful at determining the direction of blurring m_B than the spatial image.

III.2. Direct Computation of the MTF from the Blurred Image.

If the blurred image has a straight edge or line, it can be used to determine the MTF directly without recourse to oversimplified assumptions about the psf. The basic idea results from considering the blurring of a feature in the sharp image. If the feature of interest were the only feature in the image, then equations 1 through 3 would still hold with f and F receiving contributions only from the feature. The MTF would of course not change. Consequently, if noise in the blurred feature can be neglected or

suppressed then the MTF would be the ratio of the FT of the blurred feature to the FT of the suppressed sharp feature. Edges and lines are especially appealing since they are geometrically simple. For simplicity, we will limit the discussion to them. In addition it should be realized that the profile of a blurred line can be trivially obtained from the profile of a blurred edge.

In the following the edge or line of interest will be enclosed in rectangle whose sides are parallel or perpendicular to the supposed sharp line. Let vector $\vec{m}_s = [m_s, n_s]$ describe the direction and distance between pixels on the sharp line. Then the vectors

$$\vec{\imath} = [m_s, n_s] \quad \text{and} \quad \vec{\jmath} = [-n_s, m_s]$$

are respectively parallel and perpendicular to the line. Each point (m,n) located by the vector from the origin $\vec{m} = [m,n]$ is simultaneously on a j-line which is parallel to the original line and characterized by

$$j = \vec{m}\vec{\jmath} = -mn_s + nm_s$$

and or an i-line which is perpendicular to the original line and characterized by

$$i = \vec{m}\vec{\imath} = mm_s + nn_s$$

It is easily seen that $\Delta i = 1$ between neighboring perpendicular lines, $\Delta j = 1$ between neighboring parallel lines, that the distance between lines is $1/k_s = 1/\sqrt{m_s^2 + n_s^2}$ and the number of lines crossing between any adjacent points on a perpendicular line is $k_s^2 - 1$.

Neglecting end effects and assuming intensities are constant in the j-direction, then the intensity function for a sharp line, characterized by j_1 can be written as

$$I^s_{\vec{m}} = \delta_{j,j_1}$$

and that of its blurred image as

$$I_B(j) = \sum_{\vec{m}'} p_{\vec{m}-\vec{m}'} \cdot \delta_{j',j_1} \tag{6}$$

where $j = \vec{\jmath} \cdot \vec{m}$ and $j' = \vec{\jmath} \cdot \vec{m}'$.

By replacing the sums in the FT over m′ and n′ by sums over i and j, one obtains the MTF

$$P_f = \sum_j I^B(j)\, z^{\vec{k}(\vec{m}_j - \vec{m}_{j_1})} \tag{7a}$$

where the vectors m_j are to a single point on each of the j-lines. These vectors are to be selected such that the i sums in the ratio of L^b and L^s cancel. In general they are points lying closest to some line crossing the j-lines. In the case of improper focusing they can be taken to be closest to a particular i-line. In the case of motion blurring it would be those closest to a line parallel to m_s.

This problem is easily circumvented by considering the MTF on the line $\vec{k} = \kappa\vec{j}$ (κ=integer). Then

(7b) $$P_{\vec{k}} = \sum_j l^B(j)\, z^{\kappa(j-j_1)}$$

To find the MTF for improper focusing one utilizes the fact that the appropriate MTF depends only on $|\vec{k}|$ and replaces κ by $|\vec{k}|/|\vec{j}| = |\vec{k}|/\kappa$;

(7c) $$P_k^{(D)} = \sum_j l^B(j)\, z^{|\vec{k}|(j-j_1)/\kappa}$$

To find the MTF for linear motion blurring one can utilize the fact that the appropriate MTF to first approximation varies only in the $\vec{m}_B$ direction. Projection arguments easily then give

(7d) $$P_k^{(L)} = \sum_j l^B(j)\, z^{|\vec{k}\cdot\vec{m}_B|(j-j_1)/|\vec{j}\cdot\vec{m}_B|}$$

This can also be derived by writing the psf as

$$P_{\vec{k}} = \sum_q l_q\, \delta_{\vec{k},\vec{m}_q}$$

then with $l_{\vec{k}}^B = \delta_{j,k_j}$ and

$$l^B(j) = \sum_q l_q\, \delta_{j_q,\, j-j_1} \quad ,$$

$$P_k^{(L)} = \sum_q l_q\, z^{|\vec{k}\cdot\vec{m}_q|}$$

Using $\hat{m}_q \cong \hat{m}_B$, $k_j = \vec{k}\cdot\vec{m}_q \cong (j-j_1)\,\vec{k}\cdot\vec{j}/\vec{m}_B\vec{j}$ one obtains the expression for $P^{(L)}$.

It should be realized that nothing can be learned about the MTF if the motion is parallel to the edge or line. This manifests itself in the $\vec{m}_B\vec{j}$ factor in the denominator of the exponent.

IV Restoration and Noise

All digitized images are subject to noise. The noise originates from the granular nature of the original photograph, from electronic noise and round-off error in the digitization process etc. Such noise can be both signal dependent and signal independent, spatially correlated and not, it can occur as "sparkle" or spikes in the F.T. of the image. Whatever the source or nature of noise, it complicates the deblurring or deconvolution process. This is especially true in the cases under consideration, improper focusing and motion-blurring where the MTF alternates in sign. Special care must be taken to avoid emphasis of noise in regions where the MTF changes sign, i.e. along the zero-contours.

The standard approach to noise abatement is to use the MTF to construct a Wiener filter. Unfortunately such construction requires information on the noise that is not available. Also as pointed out by Castleman (2) such filters cannot handle image having large flat areas separated by sharp edges which is true of photographs of the shuttle. Consequently, a two-pronged approach to noise suppression was developed. The first involves using a two parameter high frequency component of the form

$$S^{-1} = 1 + \exp \alpha (f - f_c)$$

where $f = |k| = \sqrt{k^2 + l^2}$, f_c is a variable cut-off frequency and $\alpha^{-1} = -(N/2 - f_c)/4$. The function S thus is vanishingly small at low frequencies, ½ at $f = f_c$ and approaches unity at high frequencies.

The preliminary filter, H_p is then defined as

$$H_p^{-1} = P\,(1-S) + \gamma S$$

where γ is the second variable parameter. Since the MTF is normalized to unity at zero frequency, H^{-1}_p, starts at unity follows the MTF until $f = f_c$ where it then follows γS and approaches γS and approaches γ at high frequencies. The parameter γ allows one to emphasize or de-emphasize the high frequency region relative to the low frequency region.

This filter will clearly handle sparkle etc. noise at high frequencies, but it does not remove large values of H_p due to the vanishing of P at small and moderate frequencies. This is accomplished by introducing a third parameter, R, the maximum to minimum allowed values for the magnitude of H. Thus at each frequency point $|H_p|$ is tested and replaced by $R\,H_p/|H_p|$ if it exceeds R. This allows one to tune out or at least avoid over emphasis of noise near the zero contours of P.

V. Results

The main thrust of the research effort this summer was the designing and implementation of software to extract P from a blurred image. As a preliminary test of this software, it was applied to a Gaussian (5 x 5 pixel)-blurred image. Last summer's work showed that the 5 x 5 pixel windows on the Gaussian was of consequence only at high frequencies. Thus the assumption that the psf was azimuthally symmetric would be reasonable since most of the signal is in the low frequence range. The resulting restored image was not quite as good as that obtained last summer where the 5 x 5 window was taken into account, however the exercise clearly demonstrated that even in

the computer generated blurring information needed for constructing the MTF could be extracted from the blurred image.

V.1. Deconvolution of An Improperly Focused Image

Last summer an image was purposely blurred by defocusing the digitizing camera. Although there is no evidence in Figure 1a, the top 1/16 of the image is missing as a result of the digitization process. Such spatial-domaion truncation causes over shooting or ringing in subsequently filtered images (p.26 ref.1). Application of the newly developed software resulted in the MTF shown i Figure 1b and gave a somewhat better restored image than that obtained last year. Last year the psf was approximated by a disc of constant amplitude (4 pixels in radius) and surrounded by a fuzzy annulus 2 pixels in width (see Fig. 2a). As shown in Figure 2b, the psf derived from an edge in the blurred image (see Fig. 1a) is similar in size but has an intensity minimum in its center. Shown in Figure 3a,b are the Fourier transforms of the blurred and the restored image.

As observed last summer, the Fourier transforms are much more sensitive to the restoration process than the spatial images. It is suggestive that fine tuning on the restoration process is much easier to observe on the Fourier image than on the image itself. The expected ringing is again present in the spatial image. The rectangular form of the Fourier transformed image is apparently due to the digitization process since it occurs in all images digitized with its video digitization system.

V.2 Restoration of Two Images Blurred by Relative Motion

Two images were blurred by moving them during the digitization process. The first was quite severely blurred in that the motion extended over about thirty pixels. The second was more moderate in that the blurring was over about a dozen pixels. In both cases the same edge circled in Fig. 5a was used in extracting the psf from the blurred image. Shown in Figure 3b,c, and 4b,c are the Fourier transforms of the blurred and the restored images. The original sharp image is essentially that shown in Figure 1a. In both cases the "zero contours" are evident in the F.T. of the blurred image (see Fig. 3b and 3c). These are easy to recognize with the eye. The number of pixels to the first divided into N gives a good estimate of the blurring distance in the spacial domain. Similarly the Fourier transformed image can give a better determination of the direction of blurring than can be obtained from the spatial image. One simply uses PROFIL to define a line parallel to a maximum contour passing through k=0 and finds its slope parameters. Δk, Δl, then $m_B:n_B = \Delta l:\Delta k$, since the direction of blurring is perpendicular to extreme contours of P. This was used in Eq. 7d for determining P.

Clearly the software can be used to restore images blurred by improper focus or by linear motion blurring if the blurred image contains a straight line or edge.

VI Future Research

In last summer's report in addition to the program to construct MTF's directly from blurred images there were two other areas of potential research. The first of these

involved the development of filters to reduce noise. Although some work has bene done along this line (see Section IV), many of the suggestions remain relevant. The use of windowing or image completion is important for incomplete images digitized by the Sony TV camera. However, these problems may be circumventted by using the Eikonix system.

The main thrust of future work should be the development of a user-friendly-interactive-restorative system. The essential programs are now in place, it is simply a matter of working them into a "cookbook" that can guide the inexperienced user to construct for any blurred image of interests the appropriate MTF and restoration filter. Such a program would use a decision-true format shown in Fig. 8. Photographs of the Space Shuttle Vehicle during launch and landing would contain straight edges and would be deblurred using an MTF extracted directly from the blurred images. Photographs taken by the Space Shuttle crews of agricultural, desert and ocean features would require parameterized MTF's since they do not in general contain straight edges or lines. In such cases one would extract information needed to determine the parameters from their Fourier transforms. One would use a roster of the zeroes of the appropriate MTF, i.e. Airy or sinc function. WARP would then be used to rotate the sinc roster, and it stretch the roster until it overlaped the "zero" contours of F^B. This would generate the appropriate parameters, e.g. the radius of the Airy disc or the length and direction of the motion blurring.

The next step is to select the high frequency portion of the restoration filter. As presently envisioned this would have two parameters one a frequency, f_c, value which would be highest frequency where F^B is free of "sparkle" noise, e.g. N/6 or N/4 depending on the degree of blurring. The second parameter γ which determines the relative emphasis of the high frequency region with respect to the very low frequency region, could be set equal to one or varied slightly to produce a restored F_k where large amplitudes are mostly confined to the central region.

The final step would be to select the maximum to minimum amplitude ratio, R, in the restoration filter. This would be done by taking a value like 10 and seeing if the "zero" contours persist or have been replaced by contours of large amplitudes in the restored F_k. One would then raise or lower R until the amplutudes in the regions of the contours are about the same size as their neighbors.

After a restored F_k is obtained which is similar to ones of less blurred images one would then take its inverse Fourier transform and obtain the restored image. Other features such as "sparkle" removal could be added later if they show promise of further improving the process. This should be a user-friendly-interactive program which guides the use along until a reasonable restored image is obtained.

VII References

1. Moik, J. G., Digital Processing of Remotely Sensed Images, NASA SP-431, 1980.

2. Castleman, K. R., Digital Image Processing, ISBN 0-13-21365-7 Prentice-Hall Inc. 1979.

3. Gonzelez, Wintz, Digital Image Processing.

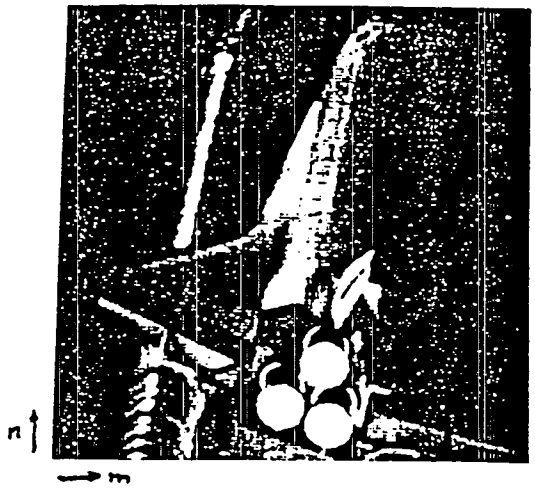
4. Press, W. H., B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, Numerical Recipes, ISB 0-521-30811-9, Cambridge Univ. Press, 1986.

5. Reynolds, G. O., J. B. DeVelis, G. B. Parrent Jr., B. J. Thompson, Physical Optics Notebook: Tutorial in Fourier Optics, ISBN 0-8194-0130-7, 1989.

6. Menke, W., Geophysical Data Analysis: Discrete Inverse Theory", ISBN 0-12-490921-3, AP, 1989.

7. Ward, R. K., B.E.A. Saleh, Restoration of images Distorted by System of Random Impulse Resonances, J. Opt. Soc. Am., A2,pp 1254-1259, 1985.

8. Tescher, A. G., H. C. Andrews, Data Compression and Enhancement of Sample Images, Applied Opts 11, #4, 919-925, 1972.



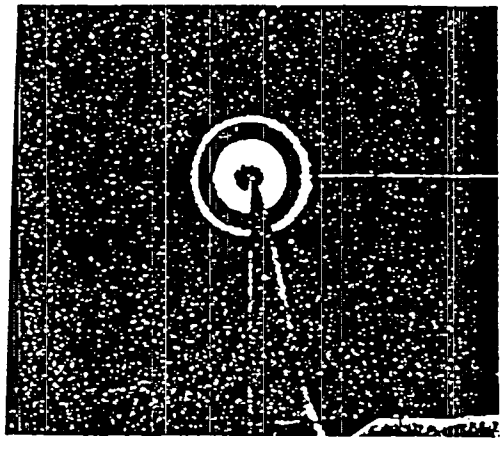
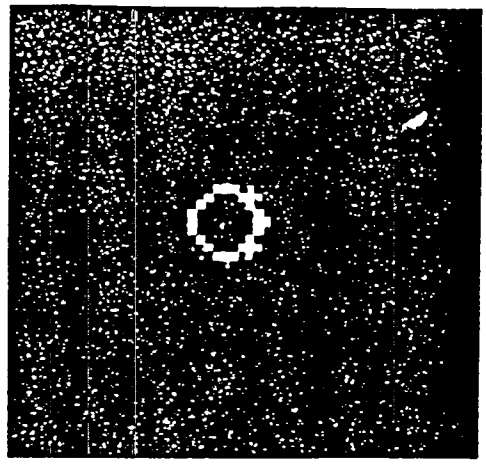


Figure 1a, b



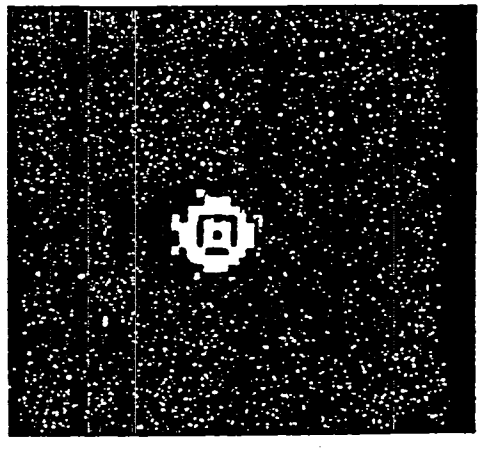
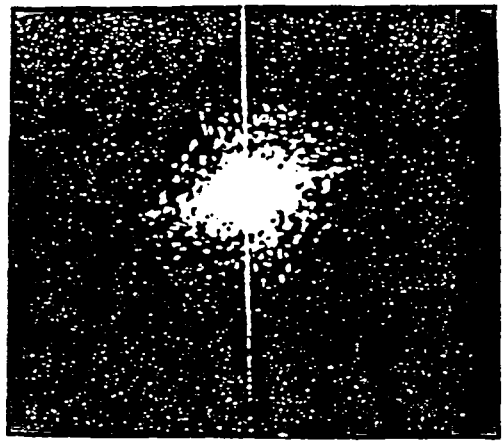
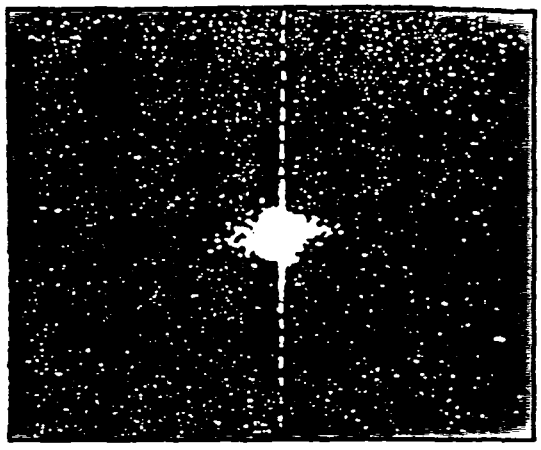
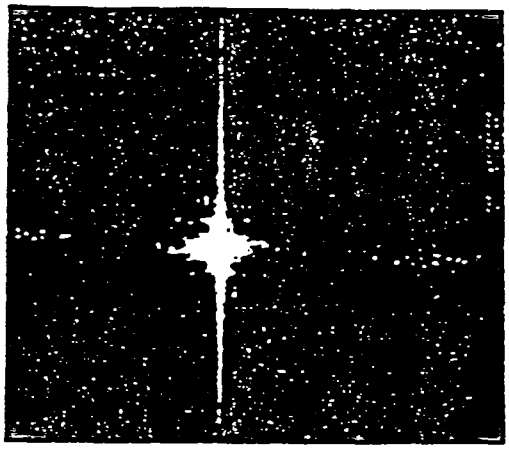
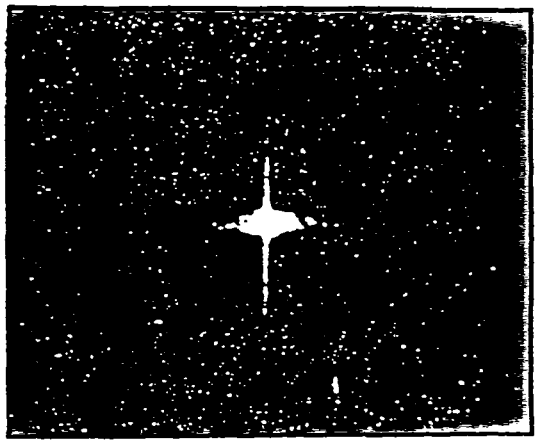


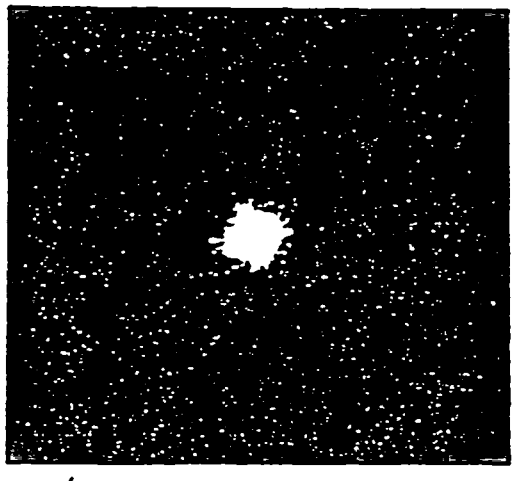
Figure 2a, b











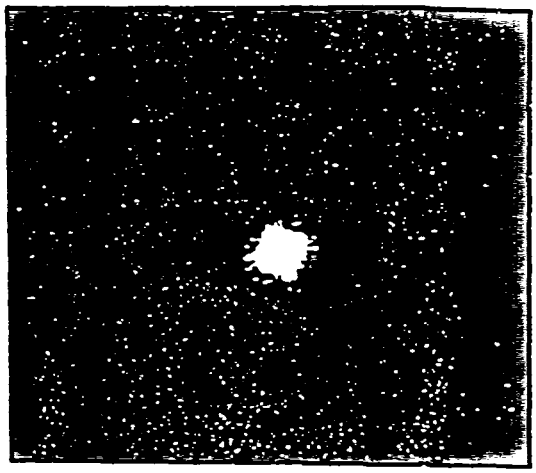


Figure 3a, b, c

Figure 4a, b, c

PHOTOGRAPHIC RESTORATION COOKBOOK

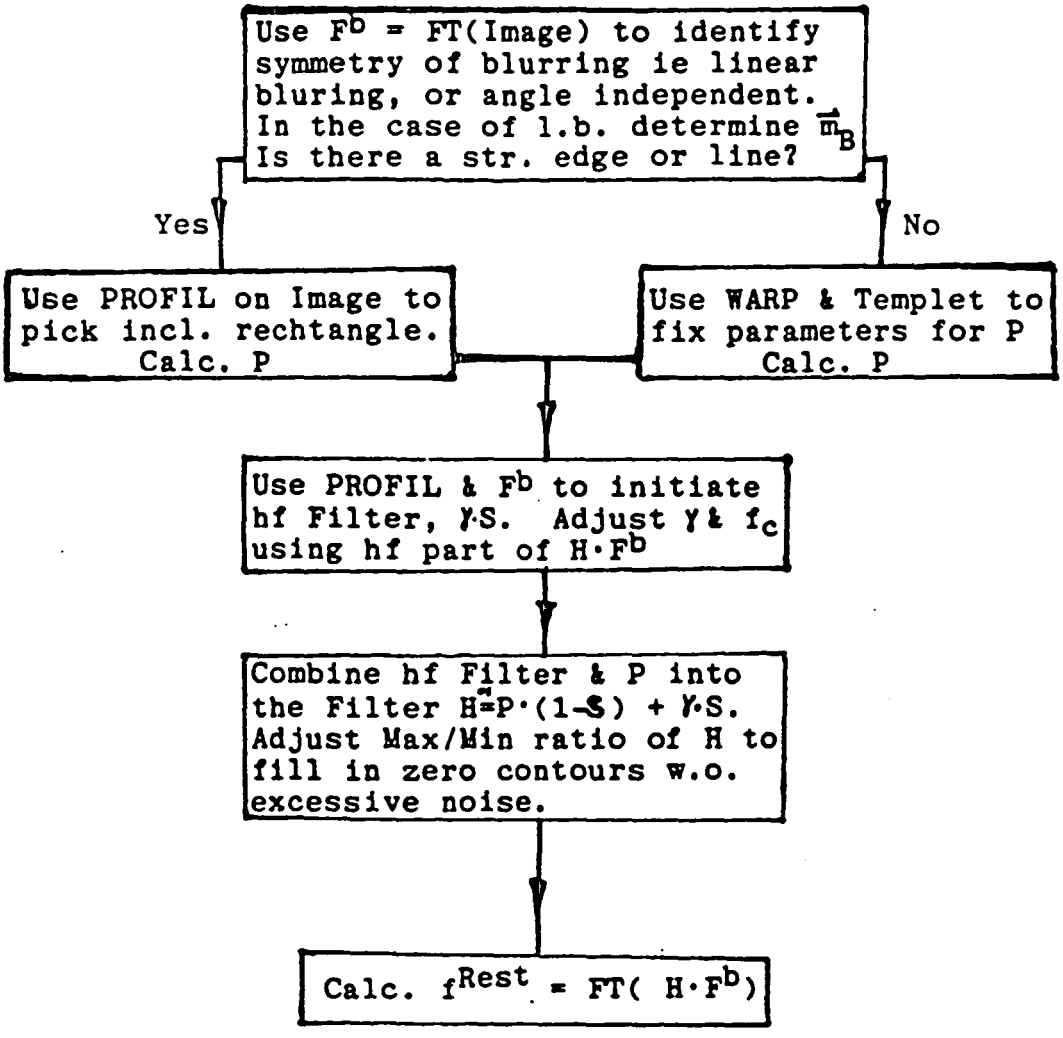


Figure 5

THE APPLICABILITY OF NONLINEAR SYSTEMS DYNAMICS CHAOS MEASURES TO CARDIOVASCULAR PHYSIOLOGY VARIABLES

Final Report

NASA/ASEE Summer Faculty Fellowship Program 1991

Johnson Space Center

Prepared By:	John C. Hooker, Ph.D.
Academic Rank:	Professor
University & Department:	Houston Baptist University Department of Biology Houston, Texas 77074

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	John B. Charles, Ph.D.
Date Submitted:	August 16, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

Three measures of nonlinear chaos--fractal dimension, Approximate Entropy (ApEn), and Lyapunov exponents--were investigated as potential measures of cardiovascular condition.

It is suggested that these measures have potential in the assessment of cardiovascular condition in environments of normal cardiovascular stress (normal gravity on the earth surface), cardiovascular deconditioning (microgravity of space), and increased cardiovascular stress (lower body negative pressure (LBNP) treatments).

INTRODUCTION

Many cardiovascular variables, such as heart rate and blood pressure, when recorded over long periods of time, appear to be random variables. Using statistics, such as the mean and the standard deviation, and measures of confidence in these statistics, statistical analyses of stochastic systems producing true random variables, are well understood and extensively used. The cardiovascular system, however, is not a stochastic system producing true random variables, but, instead, is a highly developed, very complicated, nonlinear, dynamic system producing tightly regulated output variables.

Recently, several techniques have been developed for the analysis of the behavior of deterministic nonlinear dynamic systems. These systems produce chaotic output variables that appear to be random functions of time, but which are actually completely determined by the deterministic, though nonlinear, differential equations describing the system. Even though a deterministic description of the cardiovascular system is not known, it seems appropriate to apply these techniques to its analysis.

This report presents some of the findings of a preliminary study of the applicability of three measures of nonlinear chaos to the assessment of cardiovascular condition based on recordings of cardiovascular physiology variables. These measures are fractal dimension, Lyapunov exponents, and Approximate Entropy (ApEn).

NONLINEAR CHAOS MEASURES

The three measures considered here are attempts at the quantification of the variability, or, inversely, the "patternness", of chaotic behavior. Two of these, fractal dimension and Lyapunov exponents, are used extensively in the characterization of deterministic nonlinear dynamic systems and have been used to a lesser extent in the attempt to describe the behavior of chaotic systems with unknown descriptions. The third, ApEn, has been developed for the analysis of heart rate data and, specifically, used in the analysis of fetal heart rate. A short discussion of each of these measures follows.

Fractal Dimension

Fractal dimension is a means of classifying structures that exhibit self-similarity. A self-similar structure can be described as follows. If a section of a self-similar structure is viewed in magnification, the smaller-scale section is similar in form to its larger-scale parent. As an example, consider a straight line segment which is broken into smaller segments with gaps between each segment. Then each segment is again broken into smaller segments with gaps between each of these segments, etc.. If the ratio of segment length to gap length is maintained, a self-similar structure is generated, and as each segment is viewed in magnification the same form is apparent.

For self-similar structures the fractal dimension can be thought of as a measure of how the structure fills the space of the original structure. Generally, the self-similar structure will have a fractal dimension less than the dimension of the original structure. A continuous line segment has a dimension of one, while a self-similar line structure as described above will have a fractal dimension less than one.

Consider the following larger dimensioned structures. A filled square has a dimension of two, since it completely fills a two-dimensional space. A self-similar square only partially filled with smaller self-similar squares will have a fractal dimension less than two, since the blank spaces between the squares are unfilled. Similarly, a solid cube has a dimension of three, since it completely fills a three dimensional space. A self-similar cube only partially filled with smaller self-similar cubes will have a fractal dimension less than three, since spaces between the cubes are unfilled.

For self-similar structures the fractal dimension is a function of the logarithms of the number of repeated structures and the dimension of the original structure.

Self-similar structures can be much more complicated than the examples given above. Line segments can branch into the two-dimensional plane creating a fractal dimension between one and two. Other three dimensional structures such as spheres may be nested to created fractal dimensions between two and three.

Others (1) have illustrated the appearance of self-similarity in heart rate recordings of several minutes duration. Normant and Tricot (2) have suggested an approach

to the evaluation of fractal dimension of experimentally derived curves. Their method is presently under investigation in the Cardiovascular Laboratory. It may be used to determine the fractal dimension of heart rate data and other physiological variables of interest to this lab.

Lyapunov Exponents

Lyapunov exponents are used to obtain a measure of the sensitivity of a nonlinear dynamic system to its initial conditions. Nonlinear dynamic systems are often studied in the phase space in which trajectories appear. In the case of chaotic systems these trajectories nearly repeat themselves. Lyapunov exponents are also a measure of the average rate of divergence of these trajectories. As the order of a system increases the number of Lyapunov exponents necessary to characterize the system increases.

For deterministic systems Lyapunov exponents are determined by taking the limit of the sum of the logarithms of the derivatives of the iterations of the trajectories through the phase space. This is a complicated procedure for deterministic systems, and much more difficult for systems with unknown dynamics such as the cardiovascular system. In addition, since the order of the cardiovascular system is also unknown, the number of Lyapunov exponents necessary to describe it is also unknown. However, Wolf, et.al. (3) have suggested a method of determining Lyapunov exponents from experimental time series. Their method is also under study in the Cardiovascular Laboratory.

Approximate Entropy (ApEn)

Pincus (4) has proposed another measure in the ApEn. ApEn is a measure of the deviation of a time series away from a pattern or somewhat regularity. It is a measure of the increased information available in the time series with increased variability. The algorithmic computation of ApEn is similar to the computation of Lyapunov exponents but much simpler. It involves the summation of the logarithms of the number of deviations of different magnitudes. Pincus has evaluated the level of fetal distress by determining the ApEn of fetal heart rate (4). This approach is also under consideration in the Cardiovascular Laboratory.

COMMENTS AND RECOMMENDATIONS

This preliminary study indicates that a more detailed analysis of nonlinear chaos measures applied to cardiovascular physiology variables should be vigorously pursued. There is the potential for the development of a very useful tool in the evaluation of cardiovascular condition in these measures. The analysis of chaotic systems must make use of measures other than the traditional statistics.

Present analyses indicate that these nonlinear chaos measures require time series on the order of 1000 data points in length for good resolution. For baseline data gathering situations time series of this length are easily obtained, but, for active tests such as LBNP ramp studies conducted both in the lab and in space, such time periods may not be practical. Techniques using shorter time intervals must be developed.

The following recommendations are offered:

1) The continued development of an analysis capability to be used in the evaluation of the three measures using Cardiovascular Laboratory data.

2) Using relatively long term time series of heart rate data evaluated the three measures and compare with spectral analyses of the same data. To do this and associated correlation studies an improved FFT should be incorporated.

3) Investigate the usefulness of incorporating other cardiovascular variables, such as blood pressure, into the analyses. This could have the effect of changing from a single function of time to a vector function of time, and possibly reduce the long time series requirement.

4) Consider variations of the nonlinear chaos measures to facilitate the use of shorter time series. This is of significant importance in the analysis of active LBNP ramp tests.

REFERENCES

1. Goldberger, A.L., Rigney, D.R., and West, B.J., Chaos and Fractals in Human Physiology, Scientific American, February, 1990, vol. 262:42-48.

2. Normant, F. and Tricot, C., Method for Evaluating the Fractal Dimension of Curves using Convex Hulls, Physical Review A, June, 1991, vol. 43:6518-6525.

3. Wolf, A., Swift, J.B., Swinney, H.L., and Vastano, J.A., Determining Lyapunov Exponents from a Time Series, Physica D, 1985, vol. 16:285-317.

4. Pincus, S.M., Hobbins, J.C., and Viscarello, R.P., ApEn: A Regularity Measure for Fetal Heart Beat Analysis, presented to the Society for Gynecologic Investigation, San Antonio, TX, March 20-23, 1991.

A Local Condensation Analysis Representing Two-Phase Annular Flow in Condenser/Radiator Capillary Tubes

Final Report

NASA/ASEE Summer Fellowship Program -- 1991

Johnson Space Center

Prepared by:	Amir Karimi, Ph.D.
Academic Rank:	Associate Professor
University & Department:	The University of Texas at San Antonio Division of Engineering San Antonio, Texas 78285
NASA/JSC	
Directorate:	Engineering
Division:	Crew and Thermal Systems
Branch:	Thermal Systems Analysis
JSC Colleague:	Eugene Ungar, Ph.D.
Date Submitted:	November 14, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

NASA's current effort for the thermal environmental control of the Space Station Freedom is directed towards the design, analysis, and development of an Active Thermal Control System (ATCS). A two-phase, flow-through condenser/radiator concept has been baselined, as a component of the ATCS, for the dissipation of space station thermal load into space.

The proposed condenser concept is similar to the single-phase deployable photovoltaic radiator. Here the heat rejection occurs through direct condensation of ATCS working fluid (ammonia) in the small diameter radiator tubes. Analysis of the condensation process and design of condenser tubes are based on the available two-phase flow models for the prediction of flow regimes, heat transfer and pressure drops. The prediction formulas employ the existing empirical relationships of friction factor at gas-liquid interface.

A recent preliminary analysis of condensation in the radiator tubes at Johnson Space Center revealed that annular flow covers over 90% of the tube length. This analysis uses prediction formulas for pressure drop which are, in general, based on interfacial friction factor model for steam-water flow in large tubes. The contributions made by capillary waves and transverse radius of curvature to the instability of interface are ignored in the current models.

The prediction of the interfacial friction factor and the major uncertainties associated with it play a vital role in modelling the annular flow regime. The condensation process in the proposed ATCS radiator includes two-phase flow in small tubes. The capillary waves present in the annular flow regime contribute to the liquid-vapor interfacial friction and hence have a major influence on the pressure drop and heat transfer. We have reviewed the present models of micro condensation in two-phase annular flow regime. The majority of formulations, describing the annular condensation process, are based on either on homogeneous or one dimensional separated flow models.

This study is an attempt to investigate the stability of interfacial waves in two-phase annular flow. It contains the formulation of a stability problem in cylindrical coordinates. The contribution of fluid viscosity, surface tension and transverse radius of curvature to the interfacial surface is included in this formulation. A solution is obtained for Kelvin-Helmholtz instability problem which can be used to determine the critical and "most dangerous" wavelengths for the interfacial wave. The annular condensation is examined in this study and the general equations describing the condensation process is included.

INTRODUCTION

Two-phase condensing phenomena occur in a variety of industrial applications including condenser systems associated with nuclear and conventional power plants, space power generations, thermal energy conversion systems, vapor compression refrigeration cycles, and chemical processing. The ability to understand and to model the principal physical mechanism associated with the condensing flow processes is of considerable importance.

Design and optimization of operating conditions of condensers under space environment conditions require qualitative information including pressure drop, void fraction distribution, quality distribution, flow stability. The behavior of condensation flow phenomena and the transition between flow regimes must be analyzed and understood before condenser systems may be designed for space applications. The approach adopted by the majority of the investigators has been to first focus on the changes that could occur in the flow regimes and their boundaries and also at the pressure drop at reduced-g environment. Rezkallah [1, 2] suggests that in two phase gas-liquid flow under microgravity conditions there exist a significant shift in the transition boundaries when compared with flow regime maps of the normal gravity. Under those conditions there is a pronounced influence of surface tension.

The dynamic characteristics of two-phase condensing flow have received very little attention compared to their boiling flow processes. Commonly, the flow analysis of condensation process is based on existing prediction formulas for two-phase flow models [3]. These prediction relations, in general, have been developed for two-phase, two-component systems (e.g. air-water or steam-water), which have often been obtained in rather large diameter tubes under earth's normal gravity field. The most widely used technique for determining two-phase flow regimes in tubes is that of Martinelli [4], which assumes that the pressure drop of the separate liquid and gas phases are equal, and makes no assumption regarding the specific flow configuration.

Research on condensation heat transfer in two phase flow under microgravity conditions is very limited, and experimental data for heat transfer coefficients under those conditions are sparse. Our understanding of many important physical phenomena in two phase flow remains incomplete. The microgravity, high-vacuum, low-temperature environment of space cannot be duplicated on earth , and it is too costly to place more than a few large-scale tests in space. Therefore, the heat transfer coefficient correlations for single or two-phase flow, and the pressure drop multiplier relations for two-phase flow under low gravity conditions must be determined analytically.

Surface tension forces play a significant role in determining flow regimes in small diameter tubes even at 1-g conditions. The presence of an interface between two fluid phases can influence the motion of fluids and the pressure drop when the interface has a finite curvature (transverse curvature) that is different from that at equilibrium [5]. In most studies of two-phase flow the effect of transverse curvature on the interface has been ignored and the interface has been assumed to be flat. However, the pressure drop in two-phase flow is principally due to the influence of the interfacial roughness elements. Small capillary-type waves on the interface (ripples and disturbance waves) which act as surface roughness give rise to an increased pressure drop. Since gravitational forces at 1-g are nearly dominant in such flows, surface tension has been assumed insignificant, and

therefore given little or no consideration. But, under reduced gravity conditions there is a pronounced influence of surface tension on the flow condition. These waves vary in length and in amplitude. Some amplitudes are several times greater than the mean film thickness. In two-phase flow in small diameter tubes, the interface is curved and deformed as a result of the forces acting on it. Therefore, the shape of the interface must be determined as part of the solution of the problem of two-phase flow.

The condensation process inside tubes may include several flow regimes. The most common flow regimes are annular flow (with liquid film distributed essentially uniformly along the tube wall), annular-stratified flow (with a preponderance of liquid flow along the tube bottom due to the gravitational action), wavy annular flow (in which the shearing action of higher vapor flow initiates the growth of large amplitude waves, occasionally reaching the top of the tube), slug flow (in which the regular and periodic bridging of the tube diameter results in separated slugs of vapor moving the liquid).

In a space environment there is no gravitational force acting on the liquid film. Consequently a stratified flow regime is not expected to exist. The surface wetting characteristics typical of refrigerant fluids causes uniform distribution of the condensate film along the periphery of the tube. That is, the flow is annular type. In annular condensing flow, the liquid film remains relatively thin, and has increasing velocity in the direction of flow, so that at high qualities one expects good uniform heat transfer along the entire tube periphery. The vapor velocity decreases as the vapor mass flow rate decreases more rapidly than the vapor flow area. Eventually the vapor reaches a blunt interface where all the vapor is condensed. At this point the liquid velocity is quite low since it now occupies the entire flow channel.

NASA's current effort for the thermal environmental control of the Space Station Freedom is directed towards the design, analysis, and development of an Active Thermal Control System (ATCS) to remove all the electrical loads, plus any biological or chemical generated heat loads [6]. A two-phase, flow-through condenser/radiator concept has been baselined, as a component of the ATCS, for the dissipation of the thermal load into space. Here the heat rejection occurs through direct condensation of the ATCS working fluid (ammonia) in the condenser/radiator tubes (each 10.08 ft. long with a nominal diameter of 0.09 inches). Design requirements for the condenser tubes are: 98% quality ammonia at the inlet and a minimum of 10°F subcooling at the outlet of the condenser tubes with the constraint that the pressure drop in the condenser panel should not exceed 1.4 Psid.

The pre-integrated truss heat rejection system consists of two radiator arrays: one array rejects heat from the 58°F ATCS loop and the other rejects heat from the 35°F starboard loops. Each array consists of three Orbital Replacement Unit (ORU) radiators connected in parallel. Each ORU contains eight radiator panels and has two parallel condensation paths.

A preliminary analysis of condensation in the ATCS radiator tubes at Johnson Space Center [7] revealed that annular flow covers over 90% of the tube length. The analysis were based on modified prediction formulas in [3] which contains a compilation of the existing empirical relationships for two-phase flow models for the prediction of flow regimes, heat transfer, gas-liquid interfacial friction factor pressure drops. These analysis revealed that the fluid flow is predominantly laminar in the liquid film and turbulent in the vapor core.

OBJECTIVES

The basic objective of this effort is to study the microgravity condensation process in small diameter tubes to support NASA's current mission on analysis of the two-phase flow-through condenser concept for the Space Station Freedom. The specific goals of this work to establish the forces that contribute to the friction factor at the liquid-vapor interface, to predict pressure drop and heat transfer, and to examine the interfacial instabilities contributing to the transition of flow from annular to slug flow regime. These will require a detailed analysis of the capillary waves at the liquid-vapor.

This study will be limited to examination of condensation process in the annular flow regime of the condenser tube and the examination of the capillary waves at the liquid-vapor interface and their contribution to the interfacial friction factor. In the current investigation, the modeling of the capillary waves at the liquid-vapor interface is based on potential flow theory (inviscid fluid flow). A more detailed and complete analysis requires the consideration of shear forces at the liquid-vapor interface.

We propose to use a separated two-phase flow model to study the condensing process in small diameter tubes. This requires solving the continuity, momentum and energy equations for each phase and using appropriate boundary conditions at the interface to couple the two phases together. Thus, the shape of the interface and proper formulation of tangential and normal stresses at the interface will be the focus of this research. Oscillating pressure difference component across the interface (due to surface tension and transverse curvature) which contributes to flow instability will be included in the proposed analysis.

Most studies conducted on two-phase flow condensation assume that the liquid-vapor interface is flat and ignore the influence of interfacial waves on pressure gradient. It is important to examine the behavior of the interfacial waves in two-phase flow and their contribution to pressure drop. The stability of interfacial waves is also important in determination of flow regimes in two-phase flow situations. The majority of investigations on interfacial wave behavior are based on inviscid fluid models. These works usually do not consider the effect of surface tension on interface. Most analysis of the effect of surface tension on the interface consider two-dimensional waves in Cartesian coordinate and hence do not include the contribution of the transverse radius of curvature; this is an important parameter in modeling two-phase flow in cylindrical pipes, especially in small diameter tubes.

The analytical treatment of surface waves originated with Helmholtz [8], Kelvin [9], and Rayleigh [10]. Other researchers have attempted to extend these studies. For example, Taylor focused his investigation on inertial instability of ideal homogeneous incompressible fluid. The instability of the horizontal interface between two ideal incompressible fluids dates back to G.I. Taylor [11]. The modern extension of viscid flows has been reviewed by a number of researchers who have suggested ways to include interfacial surface tension and the fluid viscosity. In recent years few studies [12,13] have considered the contribution of fluid viscosity on the " most susceptible" or "most dangerous" wavelength and the corresponding frequency.

Lienhard and Dhir [12,13] studied the viscous hydrodynamics instability of peak pool boiling problem. The formulation of their instability problem was based on a model in Cartesian coordinate, which considered a vapor jet flowing in a pool of liquid. However, in their analysis they included the contribution of transverse curvature of vapor jet to instability as a form of an additional oscillating pressure difference component across the interface. They suggested that the pressure transverse pressure effect at the liquid-vapor interface may be presented as:

$$\Delta p_{tr} = \frac{\sigma \eta}{2 R_j^2} \qquad (1)$$

where σ is the surface tension, R_j is the radius of the jet, and η (z, t) is the height of the the propagating interfacial wave.

ANALYSIS

In this work we will attempt to formulate the instability of interfacial wave in annular two-phase flow. In general, formulation of the problem will include the fluid viscosity and interfacial pressure gradient resulting from the surface tension and transverse radius of curvature. We will solve this problem for inviscid flow and make recommendations on the approach to solve the problem for viscous fluid flow.

Formulation of Stability Problem:

In formulating an interfacial wave model of annular flow we take an approach similar to that taken by Lienhard and Dhir [12,13]. However, we will use cylindrical coordinates to model our instability problem. The assumptions used in modeling the instability problem are:

1) the liquid and vapor are incompressible
2) both fluids are Newtonian
3) the base flow is steady, axisymeteric and one dimensional
4) The slope of interface is small everywhere
5) the nonlinear effects are negligible
6) the gravitational forces on the flow are negligible.

A schematic model of interfacial wave in two-phase annular flow is shown in Fig. 1. In this model vapor flows in the core of the cylinder and the liquid film flows adjacent to the wall. For this configuration we imagine a sinusoidal wave form at the interface and attempt to determine the propagation speed of the wave. The disturbance wave is assumed to oscillate as

$$\eta = \varepsilon \, e^{i k (z - c t)} \qquad (2)$$

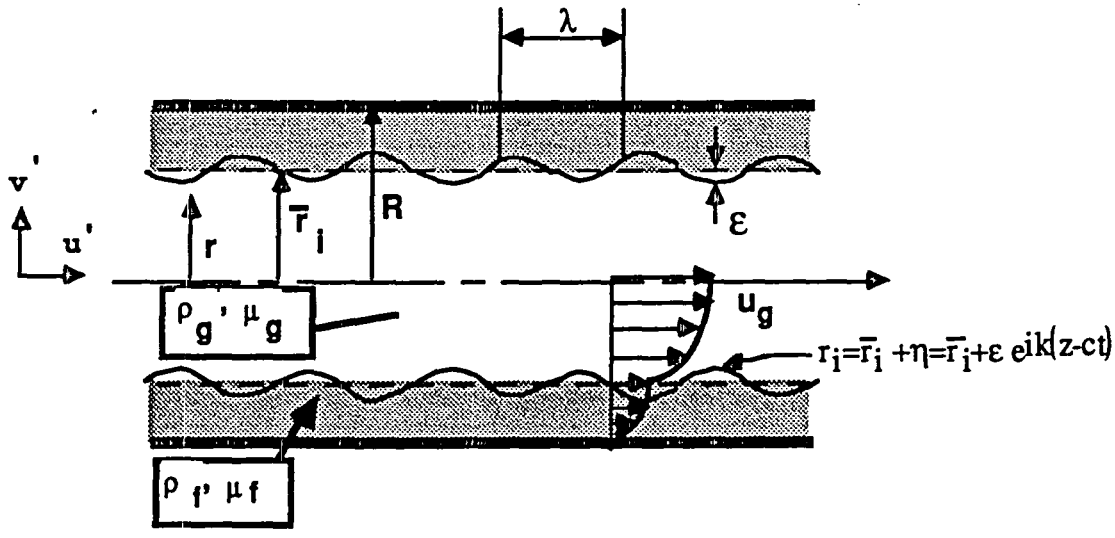


Figure 1. - Wave-shaped interface separating liquid film and vapor core.

where ε is the wave amplitude, $k = 2\pi/\lambda$ is the wave number describing the disturbance periodicity in the axial direction, λ is the wavelength $c = k\omega$ is the velocity of propagating wave and ω is the frequency of wave. The equation of interface is

$$r_i = \bar{r}_i + \eta = \bar{r}_i + \varepsilon\, e^{ik(z-ct)} \tag{3}$$

where $\bar{r}_i$ is the mean position of the interface. The criteria for the wave to be stable requires c to be real. If c is imaginary then the wave is unstable. The case $c/i < 0$ represents a decaying wave and the situation $c/i > 0$ denotes a growing wave.

Modeling the annular two-phase flow as inviscid and axisymetric, the continuity and momentum equations reduce to:

$$\frac{\partial u_i}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r v_i) = 0 \tag{4}$$

$$\frac{\partial u_i}{\partial t} + u_i\frac{\partial u_i}{\partial z} + v_i\frac{\partial u_i}{\partial r} = -\frac{1}{\rho_i}\frac{\partial p_i}{\partial z} + \nu_i\left(\frac{\partial^2 u_i}{\partial z^2} + \frac{\partial^2 u_i}{\partial r^2} + \frac{1}{r}\frac{\partial u_i}{\partial r}\right) + G_z \tag{5}$$

$$\frac{\partial v_i}{\partial t} + u_i\frac{\partial v_i}{\partial z} + v_i\frac{\partial v_i}{\partial r} = -\frac{1}{\rho_i}\frac{\partial p_i}{\partial r} + \nu_i\left(\frac{\partial^2 v_i}{\partial z^2} + \frac{\partial^2 v_i}{\partial r^2} + \frac{1}{r}\frac{\partial v_i}{\partial r} - \frac{v_i}{r^2}\right) + G_r \tag{6}$$

For the convenience and the space economy we have used the subscript "i" to represent the governing equations of both the liquid and vapor with a single set of equations (substitution of g and f for i gives relations for the vapor core liquid film region, respectively). Unless otherwise specified, we will use this subscript to represent the physical variables for both fluids with a single expression.

Now let's assume that steady annular flow $U_i(r)$ is disturbed slightly such that

$$u_i = U_i(r) + u'_i(r,z,t) \tag{7}$$

$$v_i = v'_i(r,z,t) \tag{8}$$

where the perturbed components u' and v' are infinitesimally small compared with the average velocity of the base flow (note that in condensation problem $U_i = U_i(r, z)$ and one dimensional flow assumptions might not give accurate results). The pressure can be presented as

$$p_i = p_i(z) + p'_i(r,z,t) \tag{9}$$

where

$$p' = \varepsilon\ p'(r)\ e^{i k (z - c t)} \tag{10}$$

We would like to determine whether the flow is unstable relative to the postulated disturbance. Substituting equations (7) through (10) into the governing equations (5) through (6), eliminating the steady state solution, and linearizing the results (ie. neglecting the terms of second order in u'_i and v'_i) the following equations can be written for the perturbed flow

$$\frac{\partial u'_i}{\partial z} + \frac{1}{r}\frac{\partial}{\partial r}(r\, v'_i) = 0 \tag{11}$$

$$\frac{\partial u'_i}{\partial t} + U_i\frac{\partial u'_i}{\partial z} + v'_i\frac{\partial U_i}{\partial r} = -\frac{1}{\rho_i}\frac{\partial p'_i}{\partial z} + \nu_i\left(\frac{\partial^2 u'_i}{\partial z^2} + \frac{\partial^2 u'_i}{\partial r^2} + \frac{1}{r}\frac{\partial u'_i}{\partial r}\right) \tag{12}$$

$$\frac{\partial v'_i}{\partial t} + U_i\frac{\partial v'_i}{\partial z} + = -\frac{1}{\rho_i}\frac{\partial p'_i}{\partial r} + \nu_i\left(\frac{\partial^2 v'_i}{\partial z^2} + \frac{\partial^2 v'_i}{\partial r^2} + \frac{1}{r}\frac{\partial v'_i}{\partial r} - \frac{v'_i}{r^2}\right) \tag{13}$$

These equations must satisfy appropriate boundary conditions and interfacial requirements. For the vapor core the boundary conditions are

$$v'_g = 0 \qquad \text{at} \qquad r = 0 \tag{14}$$

$$\frac{\partial u'_g}{\partial r} = 0 \qquad \text{at} \qquad r = 0 \tag{15}$$

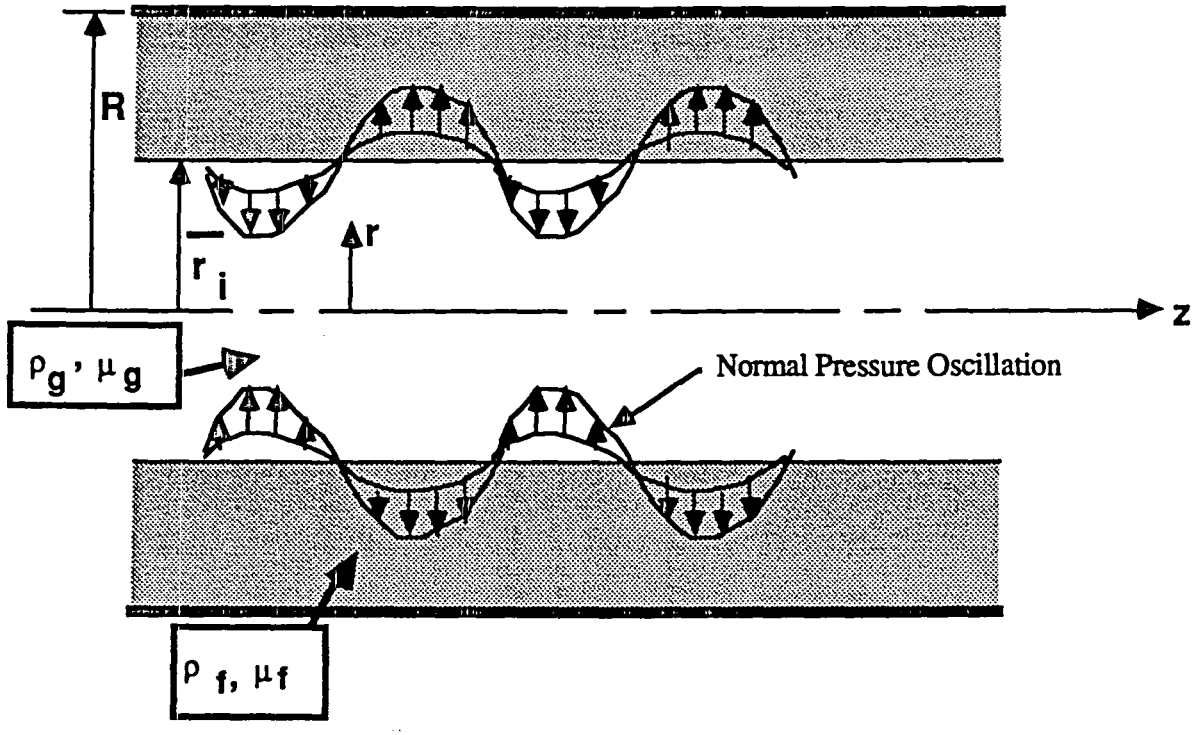


Figure 2. Pressure oscillation due to the transverse radius of the curvature of the interface.

The boundary conditions for the liquid film flow are

$$v'_f = 0 \qquad \text{at} \qquad r = R \qquad (16)$$

$$u'_f = 0 \qquad \text{at} \qquad r = R \qquad (17)$$

The following conditions are the result of force balance at the liquid-vapor interface. At the interface the normal stress exerted by the vapor is equal to those of the liquid film.

$$p_g - 2\mu_g \frac{\partial v'_g}{\partial r} + \sigma \frac{\partial^2 \eta}{\partial z^2} + \frac{\sigma \eta}{2 \bar{r}_i^2} = p_f - 2\mu_f \frac{\partial v'_f}{\partial r} \qquad \text{at} \qquad r = \bar{r}_i \qquad (18)$$

This force balance includes the surface tension contribution of pressure oscillation due to the transverse radius of the curvature of the interface as shown in Fig. 2. Note that the last term on the right hand side of equation (18) is the local time average of this pressure oscillation. A balance of tangential forces due to shear stresses of the liquid and vapor on the interface will yield

$$\mu_g\left(\frac{\partial u'_g}{\partial r}+\frac{\partial v'_g}{\partial r}\right)=\mu_f\left(\frac{\partial u'_f}{\partial r}+\frac{\partial v'_f}{\partial r}\right) \qquad \text{at} \qquad r=\overline{r_i} \qquad (19)$$

The linearized Kinematic condition at the interface requires that

$$v'_i=\frac{\partial r_i}{\partial t}+U_i\frac{\partial r_i}{\partial z} \qquad \text{at} \qquad r=\overline{r_i} \qquad (20)$$

Now we introduce the perturbation stream function as

$$\psi=\varepsilon F(r)\, e^{i k (z-c t)} \qquad (21)$$

such that

$$u'=\frac{1}{r}\frac{\partial \psi}{\partial r} \qquad \text{and} \qquad v'=-\frac{1}{r}\frac{\partial \psi}{\partial z} \qquad (22)$$

Substituting equations (21) through (22) and (10) into equations (12) and (13) we get the following relations for the equation of motion

$$\frac{1}{r}(U_i-c)\frac{\partial F_i}{\partial r}-\frac{1}{r}F\frac{\partial U_i}{\partial r}=-\frac{1}{\rho_i}p'_i+\frac{\nu_i}{i k}\left(\frac{1}{r}\frac{\partial^3 F_i}{\partial r^3}-\frac{1}{r^2}\frac{\partial^2 F_i}{\partial r^2}+\frac{1}{r^3}\frac{\partial F_i}{\partial r}-\frac{k^2}{r}\frac{\partial F_i}{\partial r}\right) \qquad (23)$$

$$\frac{k^2}{r}(U_i-c)F=-\frac{1}{\rho_i}\frac{\partial p'_i}{\partial r}-i k\nu_i\left(\frac{1}{r}\frac{\partial^2 F_i}{\partial r^2}-\frac{1}{r^2}\frac{\partial F_i}{\partial r}-\frac{k^2}{r}F_i\right) \qquad (24)$$

Eliminating p'_i by combining equations (23) and (24) the following relationship will result.

$$(U_i-c)\left(\frac{\partial^2 F_i}{\partial r^2}-\frac{1}{r}\frac{\partial F_i}{\partial r}-k^2F_i\right)+\left(\frac{1}{r}F_i\frac{\partial U_i}{\partial r}-F_i\frac{\partial^2 U_i}{\partial r^2}\right)=$$
$$\frac{\nu_i}{i k}\left(\frac{\partial^4 F_i}{\partial r^4}-\frac{2}{r}\frac{\partial^3 F_i}{\partial r^3}+\frac{3}{r^2}\frac{\partial^2 F_i}{\partial r^2}-\frac{3}{r^3}\frac{\partial F_i}{\partial r}+\frac{2k^2}{r}\frac{\partial F_i}{\partial r}-2k^2\frac{\partial^2 F_i}{\partial r^2}+k^4F\right) \qquad (25)$$

Equation (25) represents the "Orr- Sommerfeld" equation in cylindrical coordinates. Equations (23) through (25) are similar to those developed in [14]. In fact the left hand side of equations (23) through (25) are identical to the corresponding equations in [14]. However, [14] omits terms in the left hand side of those equations corresponding to equations (23) and (25).

It is more practical and useful for analysis of the instability problem to obtain a solution for equation (25) in dimensionless form . Let us define the

following dimensionless variables for the problem (for convenience we make dimensionless symbols identical to those used for the original physical variables).

$$z = \frac{z}{R}; \qquad r = \frac{r}{R}; \qquad \varepsilon = \frac{\varepsilon}{R}; \qquad \eta = \frac{\eta}{R}; \qquad \lambda = \frac{\lambda}{R}; \qquad r_i = \frac{r_i}{R}$$

$$u'_i = \frac{u'_i}{U_0}; \qquad U_i = \frac{U_i}{U_0}; \qquad v'_i = \frac{v'_i}{U_0}; \qquad v'_i = \frac{v'_i}{U_0}; \qquad p'_i = \frac{p'_i}{\rho_i U_0^2} \qquad c = \frac{c}{U_0}$$

Where U_0 is an average fluid velocity (e.g., the superficial vapor velocity). Substituting these dimensionless variables in equations (23) through (25) results in the following relations:

$$\frac{1}{r}(U_i - c)\frac{\partial F_i}{\partial r} - \frac{1}{r}F\frac{\partial U_i}{\partial r} = - p'_i + \frac{1}{i\,Re_i}\left(\frac{1}{r}\frac{\partial^3 F_i}{\partial r^3} - \frac{1}{r^2}\frac{\partial^2 F_i}{\partial r^2} + \frac{1}{r^3}\frac{\partial F_i}{\partial r} - \frac{\alpha^2}{r}\frac{\partial F_i}{\partial r}\right) \tag{26}$$

$$\frac{\alpha^2}{r}(U_i - c)F = - \frac{\partial p'_i}{\partial r} - \frac{i}{Re_i}\left(\frac{1}{r}\frac{\partial^2 F_i}{\partial r^2} - \frac{1}{r^2}\frac{\partial F_i}{\partial r} - \frac{\alpha^2}{r}F_i\right) \tag{27}$$

$$(U_i - c)\left(\frac{\partial^2 F_i}{\partial r^2} - \frac{1}{r}\frac{\partial F_i}{\partial r} - \alpha^2 F_i\right) + \left(\frac{1}{r}F_i\frac{\partial U_i}{\partial r} - F_i\frac{\partial^2 U_i}{\partial r^2}\right) =$$
$$\frac{1}{i\,\alpha\,Re_i}\left(\frac{\partial^4 F_i}{\partial r^4} - \frac{2}{r}\frac{\partial^3 F_i}{\partial r^3} + \frac{3}{r^2}\frac{\partial^2 F_i}{\partial r^2} - \frac{3}{r^3}\frac{\partial F_i}{\partial r} + \frac{2\alpha^2}{r}\frac{\partial F_i}{\partial r} - 2\alpha^2\frac{\partial^2 F_i}{\partial r^2} + \alpha^4 F\right) \tag{28}$$

where $\alpha = kR$ is the dimensionless wave number and Re_i is the Reynold number defined as

$$Re_i = \frac{U_0 R}{\nu_i} \tag{29}$$

The non-dimensional form of the disturbance wave, the equation of interface and the perturbation stream function are expressed by the following relations:

$$\eta = \varepsilon\, e^{i\alpha(z - c\,t)} \tag{30}$$

$$r_i = \overline{r_i} + \eta = \overline{r_i} + \varepsilon\, e^{i\alpha(z - c\,t)} \tag{31}$$

$$\psi = \varepsilon F(r)\, e^{i\alpha(z - c\,t)} \tag{32}$$

To analyze the stability of the interfacial waves we must solve equation (27), using relevant boundary and interfacial conditions. Note that for inviscid flow or high Reynold numbers the left hand side of the equation vanishes and it reduces to

$$(U_i - c)\left(\frac{\partial^2 F_i}{\partial r^2} - \frac{1}{r}\frac{\partial F_i}{\partial r} - \alpha^2 F_i\right) + \left(\frac{1}{r} F_i \frac{\partial U_i}{\partial r} - F_i \frac{\partial^2 U_i}{\partial r^2}\right) = 0 \qquad (33)$$

The contribution of the radial variation of the base velocity to the instability problem is represented by the last bracket in equation (33). For uniform base flow this term also vanishes and equation reduces to Helmholtz instability problem in cylindrical coordinates:

$$(U_i - c)\left(\frac{\partial^2 F_i}{\partial r^2} - \frac{1}{r}\frac{\partial F_i}{\partial r} - \alpha^2 F_i\right) = 0 \qquad (34)$$

The relevant boundary and interfacial condition for the "Orr- Sommerfeld" problem, equation (27), can be obtained from equations (14) through (20). They are

$$F_g = \frac{\partial F_g}{\partial r} = \frac{\partial^2 F_g}{\partial r^2} = 0 \qquad \text{at} \qquad r = 0 \qquad (35)$$

$$F_f = \frac{\partial F_f}{\partial r} = 0 \qquad \text{at} \qquad r = 1 \qquad (36)$$

$$p_g - \frac{2}{Re_g}\frac{\partial v'_g}{\partial r} + \frac{\sigma}{\rho_g R U_0^2}\frac{\partial^2 \eta}{\partial z^2} + \frac{\sigma \eta R}{2 \rho_g U_0^2 \bar{r}_i^2} = \left(p_g - \frac{2}{Re_g}\frac{\partial v'_g}{\partial r}\right)\left(\frac{\rho_f}{\rho_g}\right) \quad \text{at } r = \bar{r}_i \qquad (37)$$

$$\left(\frac{\partial u'_g}{\partial r} + \frac{\partial v'_g}{\partial r}\right) = \left(\frac{\partial u'_f}{\partial r} + \frac{\partial v'_f}{\partial r}\right)\left(\frac{\mu_f}{\mu_g}\right) \qquad \text{at} \qquad r = \bar{r}_i \qquad (38)$$

$$v'_i = \frac{\partial r_i}{\partial t} + U_i \frac{\partial r_i}{\partial z} \qquad \text{at} \qquad r = \bar{r}_i \qquad (39)$$

There is no complete solution available for equation (27) at this time. Most existing solutions are for the special cases of the "Orr- Sommerfeld" equation. Other solutions treat the "Orr- Sommerfeld" equation in Cartesian coordinates. The existing treatment of the instability problems are aimed at flow conditions at low Reynold number. For example, the hydrodynamic instability studies conducted by Lienhard and Dhir [12,13] are based on the formulation of problem in Cartesian coordinates. Their perturbation method solves the instability problem for a case where $Re \ll 1$ and the wave number α is close to one.

Solution of Kelvin-Helmholtz Instability Equation:

Our next attempt is to determine the most "dangerous wavelength" by solving equation (34). For the vapor core we use the boundary condition, equation (35), and the kinematic condition at the interface, equation (39). For the liquid film we employ the

boundary condition, equation (36) and the kinematic condition at the interface, equation (37).

The general solution for equation (34) is

$$F_i = C_1\, r\, I_1(\alpha r) + C_2\, r\, K_1(\alpha r) \tag{40}$$

where I_1 and K_1 are modified first order Bessel functions. Applying the boundary and kinematic conditions we obtain a solution for the vapor flow, expressed as

$$F_g = -(U_g - c)\left[\frac{r\, I_1(\alpha r)}{I_1(\alpha \bar{r}_i)}\right] \tag{41}$$

The solution for the liquid film is

$$F_f = -(U_f - c)\left[\frac{K_1(\alpha r)\, I_1(\alpha) - I_1(\alpha r)\, K_1(\alpha)}{(I_1(\alpha \bar{r}_i))(K_1(\alpha)) - (K_1(\alpha \bar{r}_i))(I_1(\alpha))}\right] \tag{42}$$

From equation (26) we get the following expression for p'_i

$$p'_i = -\frac{1}{r}(U_i - c)\frac{\partial F_i}{\partial r} \tag{43}$$

Substituting this relation into equation (37), the continuity condition of normal stress at the interface can be presented as

$$-(U_g - c)\frac{\partial F_g}{\partial r} + \frac{\sigma}{\rho_g R U_0^2}\frac{\partial^2 \eta}{\partial z^2} + \frac{\sigma \eta R}{2 \rho_g U_0^2 \bar{r}_i^2} = -(U_f - c)\left(\frac{\partial F_f}{\partial r}\right)\left(\frac{\rho_f}{\rho_g}\right) \tag{44}$$

Substituting equations (30), (41), and (42) and solving for c, equation (44) will result in a complex number, where

$$c = c_R + c_I \tag{45}$$

where c_R and c_I are the real and the imaginary part of the wave velocity . The stability condition requires that

$$c_I = 0 \tag{46}$$

The critical wave number, α_c, and wavelength, $\lambda_c = 2\pi/\alpha_c$, can be determined from this condition. The most dangerous wavelength can be evaluated using the criteria

$$\frac{\partial(-i\alpha c)}{\partial \alpha} = 0 \tag{47}$$

CONCLUSIONS AND RECOMMENDATIONS

The condensation process in annular flow has been discussed. The major forces influencing the pressure drop have been identified. It has been noted that under micro gravity conditions, the forces that influence the flow include inertial, viscous, and surface tension.

The stability of capillary waves in small tubes has been formulated. The general formulation of stability considers both viscous and surface tension effects. In addition, pressure oscillation due to the transverse radius of curvature of the interface has been included in this study. Kelvin-Helmholtz's Instability Problem has been solved. Equations (44) through (47) provide the most dangerous wavelength.

A more complete solution of equation (28) is needed to determine the viscous effect on the stability of the interface. Most models for evaluation of friction factor for condensation process in the annular flow regime are based on one-dimensional, two-phase flow in large tubes under normal gravity conditions. These models ignore the effect of capillary waves on pressure drop. These models require a friction factor multiplier based on an experimental value. All our attempts of modeling the condensation process resulted in a need for an experimental value of friction factor multiplier to determine the pressure drop due to interaction at the interface.

There is need for further study of problems considered in this work. This includes more in depth solutions to the instability problem and attempts to solve the pressure drop problem in two-dimensional flow. A more detailed and complete analysis is required to determine the influence of shear forces on the instability of the interfacial waves.

A complete analysis of the condensation problem requires solving the continuity, momentum and energy equations for each phase and using appropriate boundary conditions at the interface to couple the two phases.

REFERENCES

1. Rezkallah, K.S., "Two-Phase Flow and Heat Transfer at Reduced Gravity: A Literature Survey,"

2. Rezkallah, K.S., "Bubble-Slug Transition in Two-Phase, Liquid-Gas Flow Under Microgravity Conditions: A Preliminary Study," ASME J. Heat Transfer in Space Systems, Vol. 135, 1990, pp. 77-78.

3. Crowley, C.J., and Izenson, M.G., "Design Manual for Microgravity Two-Phase Flow and Heat Transfer (U)," Creare Inc., Final Report, Air Force Astronautic Laboratory, AL-TR-89-027/F04611-88-C-0050, Oct. 1989.

4. Martinelli, R.C., et. al., "Isothermal Pressure Drop for Two-Phase Flow in Horizontal Pipe," Trans. ASME, Feb. 1944, pp. 139-151.

5. Ostrach, S., "Low-Gravity Fluid Flows," Ann.Rev. Fluid Mech., Vol. 14, 1982, pp. 313-345.

6. Kosson, R., Brown, R., and Ungar, E. "Space Station Heat Pipe Advanced Radiator Element (SHARE) Flight Test Results and Analysis," AIAA-90-0059, 28th Aerospace Science Meeting, Jan 8-11, 1990- Reno Nevada.

7. Ungar, Eugene, "SSF ATCS Two-Phase Flow Through Radiator Preliminary Analysis - SINDA '85, FLUINT Code Development and Future Plans," NASA/JSC Internal Report, April 15, 1991.

8. Helmholtz, H. "Uber diskontinuierliche Flussigkeitsbewegungen," Monatsber. Konigigl. Preuss. Akad. Wiss. Berlin, 1968, pp. 215-228; Translation by F. Guthrie, On Discontinuous Momentum of Fluids, Philos. Mag. (4), 1868, pp. 337-346.

9. Kelvin, W., "The Influence of Wind on Waves in Water Supposed Frictionless", Proc. London Math., 1871, pp. 368-374.

10. Rayleigh, J. W. S., "On the Stability, or Instability of Certain Fluid Motions, Proc. London Math. Soc., Vol. XI, 1880, pp. 57-70.

11. Taylor, G.I., " The Instability of Liquid Surfaces When Accelerated in a Direction Perpendicular to their Plane," Proc. Roy. Soc. London, A-201, 1950, p.192.

12. Dhir K.V., "Viscous Hydrodynamic Instability Theory of the Peak and Minimum Pool Boiling Heat Fluxes," College of Engineering Bulletin, UKY BU 100, University of Kentucky, Lexington, 1972.

13. Lienhard J.H. and Dhir K.V., "Extended Viscous Hydrodynamic Theory the Peak and Minimum Pool Boiling Heat Fluxes," NASA Contract Report, NASA Cr - 2270, 1973.

14. Hewitt , G.F. and Hall-Taylor, N.S., Annular Two Phase Flow, Pergamon Press, Oxford, 1970.

RESOLVING SENSORY CONFLICT - THE EFFECT OF MUSCLE VIBRATION ON POSTURAL STABILITY

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1991
Johnson Space Center

Prepared By:	Charles S. Layne, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	Kansas State University Department of Kinesiology Manhattan, Kansas 66506
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	Millard F. Reschke, Ph.D.
Date Submitted:	August 5, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

The otolith-tilt reinterpretation hypothesis (OTTR) proposes that the CNS adapts to weightlessness by reinterpreting all otolith input as linear motion (Parker et al., 1985). This reinterpretation can be considered a strategy which the CNS uses to resolve the conflict between mismatched sensory inputs. While interpreting otolith input exclusively as linear motion is functionally useful in weightlessness it is maladaptive upon return to Earth. Astronauts have reported experiencing illusory sensations during head movement which contributes to postural instability. Disfunctional central or peripheral processes can result in sensory conflicts which, if not resolved, impair motor functioning. The purpose of this study was to assess the effect of muscle vibration in combination with a variety of sensory conflicts on postural equilibrium. The equilibrium of six healthy subjects was tested using the EquiTest sensory test protocol (NeuroCom International, INC.) with and without the confounding influence of triceps surea vibration. The sensory test conditions were randomized within two test blocks (vibration, no vibration). Sixty hz vibration was continuously applied to the triceps surea during the 20 second trials within the vibration test block. The data were analyzed with a 2x3x2 ANOVA with repeated measures with vibration, vision status, and platform status as independent variables. The subject's lowest equilibrium score from each condition was the dependent measure. Student t tests were used to assess the impact of muscle vibration within a sensory test condition. All main effects and an interaction between the presence of vision and platform sway referencing were found to be significant. Overall, a 4.5% decrease in postural stability was observed with vibration. However, equilibrium was only significantly affected in EquiTests conditions 1-4. The trend of the difference scores between conditions with and without vibration suggests that vibration is most destabilzing when the triceps surea is able to change length during postural sway (ie. conditions with a fixed support surface). The impact of sway referencing vision was virtually identical to that of eye closure, providing compelling evidence that sway referencing "nulls out" useful cues about subject sway.

INTRODUCTION

While the phenomena of postflight postural instability has not been systematically investigated, there is anecdotal evidence indicating that postural equilibrium is significantly impaired following spaceflight. For instance, returning astronauts have reported experiencing turning sensations while walking straight, perception of large pitch and rolling head movements during locomotion, and loss of postural stability when rounding corners. These illusory sensations suggest that varied sensory inputs are providing conflicting information during readaptation to the Earth's gravitational field. The inability of the perceptual-motor system to reconcile conflicting sensory inputs leads to postural disorders which will delay an emergency egress from the space shuttle.

During extended orbital flight, neural processes adapt to recalibrate the central nervous system to microgravity. Such neural adaptations result in efficient motor control in microgravity. However, the neural adaptation achieved during spaceflight is inappropriate for a 1-g environment leading to postural instability on return to Earth. The otolith tilt-translation reinterpretation (OTTR) hypothesis has been proposed to explain how the nervous system adapts to the altered vestibular input of spaceflight (Parker et al., 1985). On Earth information from the otolith receptors of the vestibular system is interpreted by the perceptual-motor system as either tilt with respect to gravity or linear motion. Since gravity stimulation is absent in microgravity, interpretation of the otolith input as tilt is meaningless. Therefore, the central nervous system adapts by reinterpreting all otolith input as linear motion. Following return to Earth and before the system readapts to unit gravity, the interpretation of all otolith input as tilt persists, producing illusions of self or environmental motion during head motion. These illusions greatly contribute of postflight postural instabilty.

The OTTR is a strategy used by the perceptual-motor system to resolve one form of sensory conflict. Specifically, the conflict between inappropriate otolith input and appropriate (for the most part) visual and proprioceptive input. A similar strategy (reinterpretation of a sensory input) may be used to resolve different types of conflict such as the conflict between

inappropriate proprioceptive input and appropriate visual and vestibular inputs.

One method by which to generate sensory conflict is through the use of muscle vibration. Muscle vibration causes rapid, alternating lengthening and shortening of the sensory region of the muscle spindle. Externally imposed rapid length changes in the sensory region distorts spindle output, causing a misperception of muscle length. Misperception of muscle length leads to inaccurate assessment of limb position, resulting in movement disorders including postural instability. Goodwin and his colleagues (Goodwin et al., 1972) were pioneers in the use of muscle vibration. They used a variety of experimental conditions to study the affect of vibration on limb position. Their results consistently indicated that the vibrated muscle was always perceived to be longer than its actual length. The perception of exaggerated length led to muscle contractions designed to shorten the muscle and return it to, what was perceived to be, the desired length. This process resulted in inappropriate limb joint angles to accomplish the intended environmental goal. Since Godwin et al., (1972) initial work, a number of investigators have confirmed their findings (Roll and Vedel, 1982; Rogers et al, 1985; Magnusson and Johansson, 1989; Pyykko, et al., 1991).

Vibration of the triceps surea (gastrocnemius and soleus) in standing subjects leads to backward sway as the muscle contracts in an effort to adjust the muscle to the perceived appropriate length. Contraction of the already appropriate length muscle leads to a new equilibrium point of the body's center of gravity (COG) posterior to the ankle joints. However, the vestibular system accurately perceives the incorrect posterior equilibrium point and attempts are made to properly realign the COG. Thus, the net affect of triceps surea vibration is increased sway relative to quiet standing as the conflicting sensory inputs "battle" each other in an effort to establish the length of the triceps surea. The unresolved conflict between competing sensory inputs during muscle vibration raises the question of the relative effects or "weighting" of unique sensory inputs and whether such weightings are context dependent.

A method by which to address the question of relative weighting of sensory inputs during postural control is to eliminate the input altogether (ex. eye closure) or manipulate the input such that it conflicts with other sensory inputs. A commercially available

postural equilibrium testing system functions to "null out" both proprioceptive inputs from the ankles and/or visual input (EquiTest system, NeuroCom International, INC.). The effect of "nulling out" proprioceptive and/or visual input on a subject's postural stability can be used to infer the extent to which a subject relies upon or "weights" a specific sensory input during standing. The generation of sensory conflicts using the EquiTest system also provides the possibility of determining if specific conflicts are resolved such that the intended goal can be accomplished. For example, visual input may be so heavily weighted during standing that conflicting input resulting from muscle vibration has no negative impact upon postural stability. Conversely, the weighting of the vibration induced proprioceptive input may be such that sway increases despite conflicting visual and vestibular inputs. It is also possible there are contextually dependent interactions of unique sensory inputs. As there are many theoretical possibilities of how a varieity of sensory conflicts are resolved, it is important to elucidate how the perceptual-motor system responds to a variety of input combinations known to produce sensory conflict. It may be that common elements exist between the various methods of conflict resolution.

The purpose of the present study was to determine the effect of triceps surea muscle vibration on postural equilibrium during a variety of conflicting sensory inputs and to assess the relative importance of vestibular, visual, and proprioceptive input during standing postural control.

METHODS

Subjects

Six individuals (three females, three males) with no diagnosed central or peripheral nervous system deficits served as volunteer subjects.

Instrumentation

In order to assess the effect of muscle vibration under a variety of conflicting sensory conditions, the EquiTest/postural testing system was used (see Introduction). The posture platform is

comprised of a potentially moving visual surround and two independently movable force plates (one for each foot) which rotate about an axis co-linear with the ankle joint. The potentially movable support surfaces and visual surround were exploited to induce a variety of conflicting sensory inputs. For example, in sensory test condition 4 the support plates were programmed to respond to the subject's anterior-posterior (A-P) sway by exactly following the degree of sway (Figure 1). In this test condition, sway about the ankle joints does not result in a change in ankle joint angle or a stretching of the ankle musculature. Thus, proprioceptive cues normally available to signal postural sway are "nulled out" with the resulting proprioception conflicting with vestibular and visual inputs. Similarly, visual inputs can be "nulled out" and therefore conflict with other sensory input. Sensory test condition 6 involves the "nulling out" of both proprioceptive and visual inputs leaving only the vestibular system to correctly indicate the true upright position with respect to gravity. A postural equilibrium score for each condition is computed by comparing the angular difference between a subject's maximum posterior and anterior COG displacements to the theoretical maximal displacement of 12.5 degrees. The score is then converted to a percentage with 100 indicating perfect stability (no sway) and 0 signaling a fall. Equilibrium scores are provided by the EquiTest system.

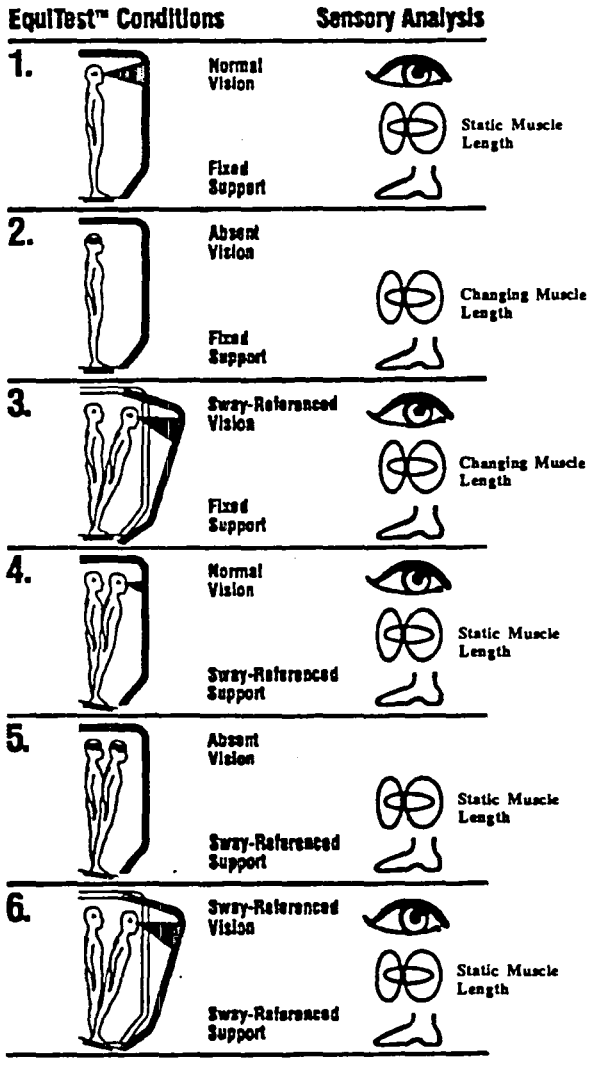


Figure 1. - Sensory test conditions

Procedures

Preliminary evidence suggested a rapid learning curve is associated with the sensory test protocol. In order to decrease the confounding influence of learning, the subjects were exposed to the six test conditions the day prior to the testing session (Day 1). Equilibrium scores were obtained during this testing session. As in all sensory test conditions, the subjects stepped onto the platform and white noise was applied through a pair of stereo headphones. This procedure is used to mask any auditory cues available to the subject from the testing environment and motion of the visual

surround. A microphone was used to communicate with the subject through the headphones. Prior to testing, the subject was fitted with a safety harness which prevents falls to the support surface.

During the testing session (Day 2), the six sensory conditions were randomly presented within blocks of the six conditions. Each condition (trial) lasted for 20 seconds and each sensory condition was presented three times for a total of 18 trials (1 block). The same procedure was followed during the sensory tests with vibration. Sixty hz vibration was applied to the lower third of the triceps surea muscle by physiotherapy vibrators held in place by Velcro bands. Vibration was applied the entire 20 seconds during sensory test conditions with vibration. Within a testing block, vibration was either always present or never present. The order the subjects received vibration (i.e. vibration in block 1 or block 2) was counterbalanced. Following completion of the first block of sensory tests the subjects rested comfortably in an adjoining room while the next subjects completed a block of tests. The first subject then returned to the laboratory to complete the second test block. This procedure prevented psychological and physiological fatigue. All tests were conducted in the Dynamic Posture Laboratory (Intermetrics, 1290 Hercules) associated with Johnson Space Center's Neuroscience Laboratories.

Statistics

In order to assess the effects of the manipulated sensory variables on postural equilibrium, the data were analyzed with 2x3x2 analysis of variance (ANOVA) repeated measures to obtain a multivariate solution. The individual subject's lowest equilibrium score in each of the conditions was the dependent measure. The first variable was vibration with the two levels being (1) present or (2) absent vibration. The second variable was vision with the three levels being (1) present, (2) absent, and (3) sway referenced. The third variable was platform status (i.e. ankle proprioception) with the two levels being (1) fixed and (2) sway referenced. This design tested whether each variable's influence on postural equilibrium was significant and whether any significant interactions between the variables influenced equilibrium. The effect of each factor was computed by subtracting the treatment mean from the baseline (no treatment) mean. The results of this procedure reflect the negative

impact of manipulating a particular sensory input on the subject's equilibrium. One-tailed Student *t* tests were used to assess the impact of vibration within a sensory condition (ex. condition 1, with and without vibration). In order to further investigate the effects of vibration under a variety of sensory conditions, difference scores were computed by subtracting the mean from a sensory condition with vibration from the mean of that condition without vibration. An alpha level of 0.05 was chosen for all statistical tests.

RESULTS

The results of the ANOVA indicate that all main effects and an interaction between vision and platform sway referencing were significant. Plotting the data revealed an interaction between platform sway referencing and the presence of vision. Table 1 lists the mean effect of each unique sensory variable.

TABLE.1- PERCENTAGES OF EFFECTS OF SENSORY VARIABLES ON POSTURAL EQUILIBRIUM

Sway Referenced Platform	-17.417%
Sway Referenced Vision	-11.875%
Absent Vision	-10.708%
Vibration	-4.472%

Vibration significantly increased postural sway in sensory conditions 1 through 4 but had no significant influence on conditions 5 and 6. There was no evidence of learning with repeated exposure to the sensory test conditions with or without vibration. Table 2 lists the mean difference in equilibrium scores between a particular sensory test condition with and without vibration. The values reflect the negative impact of vibration on the mean equilibrium score.

TABLE 2.- MEAN DIFFERENCE EQUILIBRIUM SCORES

Condition 1	-4.5*
Condition 2	-7.8*
Condition 3	-5.0*
Condition 4	-5.3*
Condition 5	-0.8
Condition 6	-3.8

*significant at .05 level

DISCUSSION

The results of the present experiment confirm that vibration of the triceps surea has a significant negative impact on postural equilibrium. This is consistent with the findings of previous investigators (Roll and Vedel, 1982; Pyykko, et al., 1991). The statistical procedures used to analyze the data resulted in the computation of percentages reflecting the negative effects of specific manipulations of sensory inputs (Table 1). The lack of significant interactions except for the presence of vision and platform sway referencing suggests that the effect of each of the sensory manipulations can be considered independent (except when the platform sway referencing and normal vision are paired) and therefore, additive. For instance, an equilibrium score of 67.403 is predicted with a sensory condition involving absent vision, vibration and platform sway referencing (-10.708 + -4.472 + -17.417 = 67.403). The actual mean was 68.000. While multiple regression analysis was not used to analyze the data (due to a violation of a specific assumption required for regression) the percentages of effect provide a good estimate of the influence of specific sensory input. While the influence of the vestibular system can not be assessed using the EquiTest system with normal subjects, the fact that normal subjects do not fall despite a variety of sensory conflicts confirms previous reports that vestibular input is the

dominant arbitrator of postural control (Forssberg and Nashner, 1982).

The small influence of vibration on postural stability, though significant, is not surprising. Vibration primarily affects muscle spindle firing patterns which are only one of many inputs contributing to proprioception. The percentage of effect observed when the platform is sway referenced provides an estimate of the impact of all proprioceptive inputs on postural stability. As expected, the combined influence of joint receptors, muscle spindles, golgi tendon organs and pressure receptors is far greater than that of muscle spindle input alone.

The present data provides compelling evidence that sway referenced visual input has the same effect on postural stability as absent vision does (Table 1). Thus, sway referenced vision neither improves or decreases equilibrium relative to absent vision. This supports the NeuroCom system manufactures' claim that vision is being "nulled out" in the sensory test conditions involving sway referenced vision. It also suggests that the strategy used by the perceptual-motor system to resolve the conflict between inappropriate visual input and other sensory inputs is to ignore the visual input. A simple way to think about the effects of sway referencing vision is that it "blinds" the postural control system to visual inputs. Conversely, a blind individual, free of vestibular impairments, should perform equally as well as a sighted individual experiencing sway referenced vision.

While the negative impact of vibration averaged about 4.5 % across conditions, all sensory test conditions were not equally effected. The difference scores in Table 2 reflect the fact that only sensory test conditions 1-4 were significantly influenced by vibration. The question of why vibration differentially impacts postural equilibrium under different sensory conditions remains to be addressed. Vibration affects both the primary and secondary muscle spindle endings. Primary endings preferentially respond to changes in muscle length while secondary endings are primarily sensitive to velocity changes of the spindle's sensory region. Equilibrium testing using a fixed support surface will result in changes in muscle length at certain velocities in response to subject sway. Thus, both primary and secondary spindle endings will be activated. Vibrating a muscle of changing length will create a conflict between the sensory input associated with the length

change and the vibration induced input. Alternatively, postural sway with a sway referenced support surface will have minimal impact on muscle spindle firing characteristics since subject sway does not alter muscle length. During platform sway referencing the vibration induced input will not interact with input signalling changes in muscle length. Therefore, sensory tests conditions utilizing platform sway referencing offer an opportunity to assess the independent effects of muscle vibration.

Interaction Between Changing Muscle Length and Vibration

The following section offers hypotheses about the possible interaction between changing muscle length and vibration. These hypotheses are based on the trend of the mean difference scores between selected conditions. None of the differences in Table 2 were signficant though the difference between conditions 2 and 5 approached significance (p=.08).

Sensory condition 2 provides the combination of inputs during which vibration is predicted to have the greatest effect on postural stability. Sensory condition 5 provides a sensory input combination on which vibration is predicted to have minimal impact on equilibrium. In both conditions 2 and 5 vision is absent. What is different is that in condition 2 the triceps surea changes length during postural sway. As hypothesized, the influence of vibration is increased relative to condition 5 (Table 2). This finding supports the idea of an interaction between vibration and changing muscle length. The minimal impact of vibration in condition 5 suggests that proprioceptive input associated with subject sway is effectively "nulled out" when the platform is sway referenced.

Given the effects of vision sway referencing are nearly identical to those of eye closure (Table 1), and that platform sway referencing "nulls out" proprioceptive input, the results from sensory test conditions 5 and 6 should be similarly effected by vibration. While the difference score in condition 6 is 3 percent greater than condition 5, these conditions were the only conditions not significantly effected by vibration (Table 2). The difference scores for conditions 5 and 6 also support the idea that applying vibration to a muscle of static length will have limited impact. Sensory test condition 3 involves sway referenced vision with subject sway resulting in muscle length changes. Thus, conditions 2

and 3 can be considered equivalent and the difference scores reflect this equivalence.

The evidence presented up to now suggests that vibration strongly contributes to postural instability when imposed on muscles of changing length. The effect of applying vibration in sensory test condition 4 presents a paradox. Since condition 4 involves platform sway referencing, applying vibration would not be predicted to have a significant impact on postural equilibrium. However, the data do not support such a conclusion. Although the platform is not sway referenced in condition 1, it can be argued that there are only small changes in muscle length due to the lack of subject sway, therefore sensory test conditions 1 and 4 can be considered equivalent. Since both of these conditions involve normal vision it appears likely that vision interacts with vibration to decrease postural stability despite the lack of changes in muscle length. While it is recognized that this final section is highly speculative it does provide a starting point for future investigations.

CONCLUSIONS

The results of the present study support the following conclusions:

1) The effects of specific sensory input manipulations on postural stability can be fairly accurately quantified.

2) The NeuroCom postural stability testing system effectively "nulls out" both vision and proprioceptive inputs associated with postural sway.

3) The negative impact of triceps surea vibration on postural stability is greatest when imposed during sensory test conditions allowing changes in muscle length.

REFERENCES

Forssberg, H. and Nashner, L.M. Ontogenetic development of postural control in man: adaptation to altered support and visual conditions during stance. The Journal of Neuroscience. 2:545-552, 1982.

Goodwin, G.M., McCloskey, D.I. and Matthews, P.B.C. The contribution of muscle afferents to kinaesthesia shown by vibration induced illusions of movement and by effects of paralysing joint afferents, Brain, 95:705-748, 1972.

Magnusson, M. and Johansson, R. Dynamic performance of vibration induced anterior-posterior sway during upright posture in normal subjects. Acta Otolaryngol., Supplement 468: 227-230, 1989.

Parker, D.E., Reschke, M.F., Arrott, A.P., Homick, J.L., and Lichtenberg, J.L. Otolith tilt-translation reinterpretation following prolonged weightlessness: Implications for preflight training. Aviation, Space and Environmental Medicine. 56: 601-606, 1985.

Pyykko, I., Enbom, H., Magnusson, M., and Schalen, L. Effect of proprioceptor stimulation on postural stability in patients with peripheral or central vestibular lesion. Acta Otolaryngol, 111:27-35, 1991.

Rogers, D.K, Bendrups, A.P., and Lewis, M.M. Disturbed proprioception following a period of muscle vibration in humans. Neuroscience Letters, 57:147-152, 1985.

Roll, J.P. and Vedel, J.P. Kinaesthetic role of muscle afferents in man: studies by tendon vibration and microneurography. Experimental Brain Research., 47:177-190, 1982.

TRAINING, QUALITY ASSURANCE FACTORS, AND TOOLS INVESTIGATION: A WORK REPORT AND SUGGESTIONS ON SOFTWARE QUALITY ASSURANCE

Final Report

NASA/ASEE Summer Faculty Fellowship Program–1991

Johnson Space Center

Prepared By:	Pen–Nan Lee, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	University of Houston – University Park Department of Computer Science Houston, Texas 77204–3475

NASA/JSC

Directorate:	Safety, Reliability and Quality Assurance Office
Division:	Quality Assurance and Engineering Division
Branch:	Software Product Assurance Office
JSC Colleague:	Vincent D. Watkins
Date Submitted:	August 16, 1991
Contract Number:	NGT–44–001–800

ABSTRACT

In the past Summer, several research tasks have been conducted, some observations were obtained, and several possible suggestions have been contemplated involving software quality assurance engineering in JSC. This report briefly describes these research tasks. Next, it gives a brief discussion on the role of software quality assurance in software engineering, following that are some observations and suggestions. A brief discussion on a training program for software quality assurance engineers is provided. A list of assurance factors as well as quality factors are also included. Finally, a process model which can be used for searching and collecting software quality assurance tools is presented.

1. INTRODUCTION

It has been a very pleasant and enjoyful experience to work with NASA/JSC software quality assurance engineers during this past Summer. The JSC Summer faculty fellow program office has provided us with a great deal of assistance including providing weekly seminars and tours, special events and activities, such as the SOAR'91 conference, and arranging meetings with JSC management. Every Summer faculty fellow also received the necessary help and convinence within each individual branch to make his/her task work done in the best possible way.

In the past 10 weeks, I worked on several tasks for Quality Assurance Engineering Office. One task was to search and collect materials and references related to software quality assurance engineering, and JSC technical library's RECON data base was the primary resource. Several thousand articles on quality and software engineering were saved to diskette. Key words used in the search for the quality articles were determined by searching JSC technical library's ARIN data base for general books on quality. The software life cycle was used to determine key words to use for searching for articles on software engineering. A list of 688 references in 90 pages has been produced. However, there are more references left to be collected.

The other task was to investigate tools which can be utilized by SQA engineers. After weeks of contacting dozens of organizations, I found that the best possible way could be to coordinate with JSC's Software Technology Brance (STB) and MITRE for the matter. They have a great volume of software available and free, unless otherwise noted, to needed organizations. Assistance in identifying, evaluating and developing software technologies and tools can also be attained from them. I accessed the "library X", but found none of the about 180 software programs has merits for direct usages in software quality assurance. However, there is more software, such as Autolib, COSTMODL, COMPASS, REAP, etc., which deserve a chance to be appraised and determined based on SQA needs.

I also tried to observe the role of SQA in JSC's software engineering and to give suggestions. A preliminary impression is obtained and included in this report. The rest of this report gives some details of my observations and suggestions.

2. THE ROLE OF QUALITY ASSURANCE IN SOFTWARE ENGINEERING

As mentioned in the introduction, only an impression was obtained due to the short period of time. Before expressing my observation, some issues regarding SQA in software engineering in general deserve a brief discussion. Within the context of "role", two issues are briefly stated in terms of quality assurance personnel.

2a. WITHIN THE ORGANIZATION

There are controversial opinions regarding the necessity of having an independent software quality assurance (SQA) group. Arguments for opposing an independent SQA include wasting resources or else other key problem areas could have made the endeavor a greater success, increasing costs, to nurture counterproductive attitudes such as they-against-us, not-our-responsibility, and the total quality management is not possible.

On the other hand, however, a good independent SQA group can shake that complacency of the development organization. The SQA group should become involved before a project begins and work with the managers and developers throughout the life of a project. As a team player, a good independent SQA and total quality management are complementary. In order to promote SQA staff as part of the team and achieve their trustworthiness, effective methods and tools have to be in place and necessary training must be available.

Managers and developers may also be provided with these methods and tools if necessary, hence, the total quality management is emplaced and high quality software is guaranteed.

2b. BETWEEN THE CONTRACTING ORGANIZATION AND ITS CONTRACTORS

Three different roles could be adopted, and each role has its own importance. The role of an "approver" is usually to make, or to make jointly, the decision of accepting and thus allowing the software to be used, and hence, has the decision authority as well as responsibility. The other role is the "evaluator" who's task is to attain a high level of understanding of the system, and to make a judgement as to the reliability based on criteria indicating the assurance level required. The criteria and assurance levels are set by the "policy authority", the third possible role, who is able to issue a certificate based on evaluation. For statutory integrity requirements, these three roles should not be mixed.

The software procurement process should include a set of assurance metrics or measures which would have to be applicable throughout the entire procurement process:

1. Procurement of a major system shall start with a definition of requirement and a risk assessment for the system. Based on the risk assessment, the requisite assurance level for each aspect of the system software should be specified.

2. The measures must be applicable during software development in order that the development process can be geared to produce an engineered level of assurance for the software.

3. The measures should be capable of use by an evaluation team. Specifically, they should form the basis of the assurance judgements made in evaluaton.

4. If a component is to be used in a number of applications, it is necessary to have a "certificate" indicating the assurance level of the component independent of its application. One would expect the assurance measures to be usable in certification.

5. The measures must help in making the decision whether or not to accept a given software in a particular environment. One should be able to compare the achieved assurance levels with the required assurance levels and determine whether or not the implemented software meets or exceeds its requirements.

2c. OBSERVATION

My impression is that SQA engineers in JSC are basically overloaded with work and duty, i.e. under-staffed. If the situation continues, two obvious things will happen: 1) a SQA staff member can not perform the best way he/she could due to lack of time, and 2) he/she must assume more duty and that may blur his/her role. One way to resolve this is to recruit more SQA engineers and provide them with quality training. The concept, which is under development in JSC, of attaching a software quality expert to each project throughout its life-cycle is an effective idea. As mentioned earlier, the expert will be part of the team, and he/she can handle the quality issues in details and in daily bases.

3. DIRECTIONS AND ADVANCEMENTS

Three discussions are given in the areas of training, assurance and quality factors, and investigation of tools. In section 3a, developing and delivering a joint training program with a major local university and industrial software developers is discussed. Section 3b discusses the assurance and quality factors, requirements of assurance measures, and the need of metrics. In section 3c, a process model for investigating SQA tools is presented

3a. TRAINING PROGRAM

It is essential to have skilled software engineering and quality assurance engineering work forces within NASA/JSC as well as its contractor companies. However, highly qualified software engineers are normally attained through training and extensive experience. The main reasons for the lack of these engineers are that most universities do not prepare students to develop industrial software and little is available in the way of continuing professional development.

It is important to develop and deliver software engineering and quality engineering training. By cooperating with a major local university and collaborating with key software developers, a professional training center could be established. Both dedicated consultant staff as well as full time professional training staff could work on developing the training program, to supply the knowledge and to teach the detailed techniques to the trainee.

An organization should establish a policy which requires every employee to receive a minimum number of hours of training per year. These hours, for examples 30 or 40, should be determined based on the organization's scope of needs.

The training center should provide various levels of training programs for a variety of participants, such as senior software engineers, project managers, etc. A basic program may contain a list of fundamental courses, such as Concepts in Programming, Ada, Operating Systems, etc. An advanced training program may contain courses in areas such as:

Project Management
Software Design
Configuration Management

Testing
Metrics
Risk Assessment
Software Process Development and Improvemen
Contracting Issues
etc.

3b. ASSURANCE FACTORS AND QUALITY FACTORS

Technical staff should be concerned with assurance throughout the entire procurement process as well as the installed system in the operational environment. They should be able to and must attain a set of "sound" assurance metrics for some given application domain. The soundness of the measures sould be assessed through the normal scientific proven process, or, in practice, employ the measures, after peer review and reasonable confidence are attained, in real world situations in order to provide evidence of how they work in practice.

The requirements of measures are concerned with the basis which is essentially a comparative analysis of the capabilities of different approaches to software development, and which identifies a range of factors which contribute to assurance. The following is a preliminary set of requirements for measures which should be helpful in any review process and which drive the selection of the technical factors which are believed contribute to assurance.

1. Scale: The measures should be applicable across the spectrum of computer systems.

2. Assurance Levels: The range spans from that gained in well tried commercial software to the highest level achievable with current software engineering techniques.

3. Uses: Applicable measures in specifying requirements, to assist in performing evaluation, to be usable as guidelines in development, and to form a basis for certification.

4. Users of the Measures: Meaningful measures to both "laymen" and technical staff in developing and assessing systems.

5. Consistency in Application: Insurance that consistent results are produced by different evaluation teams.

6. Profiles: A derivable assurance profile for a system, i.e. to assess the assurance in individual functions and properties of a system, rather than giving a blanket assurance level for the whole system:
   a) cost effective to have an assurance profile where there are high assurance defenses against the most serious threats, but where the rest of the system is developed to a lower level of assurance, than to develop a complete system to a uniformly high assurance level,

b) not desirable to have to give a low assurance level to a complete system if a single flaw is found which only affects one minor function. Thus, it is desirable both to specify an assurance profile in requirements, and to produce a profile in evaluation.

Factors contributing to assurance are essential to the applications. The following is a preliminary list of assurance factors which cover facets of the procurement process from initial elicitation of requirements through to the procedures for installing and updating the software in the operational environment.

1. Development: The concern with the technology deployed in the production of the software components.

2. Requirements: The confidence in the accuracy of the model of the operational environment for the software, and the normal software engineering issues of confidence that the requirements elicitation process has correctly identified the specific requirements for the system.

3. Architecture: A high-level design description which identifies the trusted and untrusted components and the interdependencies between the components.

4. Evaluation: The proper checking of the development process, and analyzing and testing the product to look for flaws.

5. Configuration Control: The concern with the ability to identify all components of the system, through all stages of development and evauation, and to control how they are changed.

6. Complexity: A major factor which impacts comprehensibility; the more complex the less assurance.

7. Human Interface: The confidence that the users would be able to comprehend the information from the computer system even during a safety-critical event.

8. Staff Issues: Related to the skills of the staff employed in development and evaluation.

9. Tools: Assurance in the tools which are used in support of development and evaluation.

Another key issue is to attain a set of quality measures. As mentioned above, the set of measures must be sound, that is fewer failures occur in systems judged to have high assurance than those judged to have low assurance. The difficulty is that there is no agreed objective measure of relevant aspects of software such as

reliability. The following is a preliminary list of quality factors. The need is to design measurable requirements.

1. Correctness: deals with the extent to which the software design and implementation conform to the stated requirements.

2. Efficiency: deals with the resources needed to provide the required funtionality.

3. Expandability: deals with the perfective aspects of software maintenance, i.e. increasing the software's functionality or performance to meet new needs.

4. Flexibility: deals with the adaptive aspects of software mantenance, i.e. modifying the software to work in different environments.

5. Integrity: deals with security against either overt or covert access to the programs or data bases.

6. Interoperability: deals with how easy it is to couple the software with software in other systems or applications.

7. Maintainability: deals with the ease of finding and fixing errors.

8. Manageability: deals with the administrative aspects of modification to the software.

9. Portability: deals with transporting the software to execute on a host processor or operating system different from the one for which it was designed.

10. Usability: deals with the initial effort required to learn, and the recurring effort to use, the functionality of the software.

11. Reliability: deals with the rate of failures in the software that render it unusable.

12. Reusability: deals with the use of portions of the software for other applications.

13. Safety: deals with the absence of unsafe software conditions.

14. Survivability: deals with the continuity of reliable software execution in the presence of a system failure.

15. Verifiability: deals with how easy it is to verify that the software is working correctly.

3c. A PROCESS MODEL FOR SQA TOOLS INVESTIGATION

The following is a process model for searching SQA tools. The model can also be used as a basis, i.e. some modifications may be needed, for coordinating with STB and MITRE for searching SQA tools. As mentioned earlier, both STB and MITRE can assist in performing the search task. This model contains 9 phases. The first three phases may not be necessary for someone who has been involved in JSC's SQA activites for long time, while the other six phases are needed in general:

PHASE

1. perform preliminary needs analysis
2. characterize the existing culture
3. analyze and summarize results of the characterizations: preliminary problem identification
4. identify candidate tools/environments that can meet requirements or improvement needs
5. present findings and obtain feedback
6. present final recommendations and plan for evaluation
7. arrange for evaluators from JSC/contractors to participate in decision
8. develop transition and insertion plan
9. develop comprehensive training plan

4. CONCLUSION AND ACKNOWLEDGEMENTS

It has been a very pleasant and enjoyful experience to work with NASA/JSC software quality assurance engineers. In the past Summer, several research tasks have been conducted, some observations were obtained, and several possible suggestions have been contemplated. This report describes these tasks, observations and suggestions.

I would like to thank E. Joseph Ripma and Vincent D. Watkins of the Software Product Assurance Office for their support and patience, George Neil of UNISYS for his enthusiastic attitude on software procurement, Ernest Fridge of STB and Dona Erb of MITRE for their assistance in searching for SQA tools and help in forming the process model mentioned in section 3c, and finally, Robert Youngblood and Bernice Mays of LORAL for sharing their office with me.

5. REFERENCES

BUCK89 Buckley, F., *Implementating Software Engineering Practices* John Wiley & Sons, New York, 1989.

CHAR89 Charette, R. N., *Software Engineering Risk Analysis and Management*, McGraw–Hill Book Company, New York, 1989.

DEIM90 Deimel, L. E., editor, *Software Engineering Education, SEI Conference 1990*, Springer–Verlag, New York, 1990.

DEUT88 Deutsch, M. S. and R. R. Willis, *Software Quality Engineering*, Prentice Hall, Englewood Cliffs, New Jersey, 1988.

DUNN90 Dunn, R. H., *Software Quality: Concepts and Plans*, Prentice Hall, Englewood Cliffs, New Jersey, 1990.

EVAN89 Evans, M. W., *The Software Factory*, John Wiley & Sons, New York, 1989.

GERA86 Gerard, M. and P. W. Edwards, *Strategies for Revitalizing Organizations, Proceedings of the Second NASA Symposium on Quality and Productivity*, Washington, D.C., December 2-3, 1986.

HUMP88 Humphrey, W. S., "Characterizing the Software Process", *IEEE Software*, Vol. 5, No. 2, March 1988, pp. 73-79.

NASA91 NASA, *SMAP DIDS - Information System Life-Cycle and and Documentation Standards, Version 4.3., 1991*

PERR91 Perry, D. E., and G. E. Kaiser, "Models of Software Development Environments, *IEEE Transactions on Software Engineering*, Vol. 17, No. 3, March 1991, PP. 283–295.

PRES88 Pressman, R. S., *Making Software Engineering Happen,* Prentice Hall, Englewood Cliffs, New Jersey, 1989.

SENN89 Sennett, C. T., editor, *High–Integrity Software*, Plenum Press, New Tork, 1989.

VINC88 Vincent, J., Waters, A. and J. Sinclair, *Software Quality Assurance*, Volume I, & II, Prentice Hall, Englewood Cliffs, New Jersey, 1988.

WALT91 Walters, N. L., "An Ada Object-Based Analysis and Design Approach", *ACM Ada Letters*, Vol. 11, No. 5, SIGAda, ACM, July 1991.

PLANNING AND RESOURCE MANAGEMENT IN AN INTELLIGENT AUTOMATED POWER MANAGEMENT SYSTEM

Final Report

NASA/ASEE Summer Faculty Fellowship Program -- 1991

Johnson Space Center

Prepared By:	Robert A. Morris, Ph.D.
Academic Rank:	Assistant Professor
University and Department:	Florida Institute of Technology Computer Science Dept. Melbourne, Florida 32901
NASA/JSC Directorate:	Engineering
Division:	Propulsion and Power
Branch:	Power
JSC Colleague:	Thomas D. Jeffcoat
Date Submitted:	August 23, 1991
Contract Number:	NGT-44-001-800

ABSTRACT

Power system management is a process of guiding a power system towards the objective of the continuous supply of electrical power to a set of loads. Spacecraft power system management, unlike most ground-based procedures, requires planning and scheduling, since electrical power is typically a scarce resource in space. This fact will be most evident in future longer-term space missions such as Space Station Freedom, lunar and Mars bases. The automation of power system management for future spacecraft has been recognized as an important research and development goal. Current automation technology has evolved to include microprocessing capabilities in switchgear, applying data and numerical processing capabilities to the retrieval, storage, and analysis of sensor and switchgear data, and the use of expert systems for automating human problem-solving capabilities. As examples of the latter, several rule-based expert systems for fault diagnosis and load scheduling have emerged. However, no single expert system capable of performing all the functions of power management has yet been developed. Presumably, the memory and computational demands exceeds the capabilities of a single, centralized expert system. In addition, it is highly unlikely that such a system could meet all the requirements for effective power system management, which includes a very fast response time for fault diagnosis and recovery. Such requirements are especially crucial in a hostile environment such as space, with limited redundancy capabilities in power system components. It is questionable, therefore, whether current-generation expert system technology is applicable for power system management for space. The objective of the ADEPTS project at JSC is to investigate new techniques for power management automation. These techniques involve integrating current expert system technology with that of parallel and distributed computing, as well as a distributed, object-oriented approach to software design. An overall design of the ADEPTS power management system was proposed as the result of the author's first summer fellowship [7]. The work performed during the second term, which is summarized in this report, is a further contribution to the ADEPTS development effort. The focus of the current effort is the integration of new procedures for automatically planning and scheduling loads with procedures for performing fault diagnosis and control, already implemented in ADEPTS. The objective is the concurrent execution of both sets of tasks on separate transputer processors, thus adding parallelism to the overall management process.

INTRODUCTION

The objective of the activity summarized in this report is the integration of planning and resource management into the ADEPTS power management system. Planning activity takes the form of scheduling, rescheduling and executing loads, and the automation of this activity should admit interaction with the human planner. Resource management consists of monitoring the supply of power to the overall system, configuration management, and load shedding.

The immediate purpose of this research activity was two-fold: first, to add more autonomy to ADEPTS, where by autonomy is meant the ability to perform control functions without the intervention of human control, and secondly, to further investigate the use of parallel processing for the performance of power management for spacecraft.

In addition to developing these ADEPTS enhancements, a more detailed overall design of a more robust ADEPTS system was constructed during this research activity. This work furthered the effort of last year's summer term (summarized in [7]), and is also summarized in this report.

ADEPTS PLATFORM

Introduction to ADEPTS

ADEPTS (Advanced Electrical Power Management Techniques for Space systems) is an example of a "reactive, real time system" [4, p. 997]. By reactive is meant a non-terminating program which maintains an on-going interaction with its environment. By real time is meant the imposition of timing constraints on the execution of the tasks of the system. [1, p. 1]. The environment here comprises both the power system itself, as well as the human operator (crew member or mission control personnel).

ADEPTS uses distributed memory and distributed processing for power system management. Independent, communicating processes cooperate to solve problems related to diagnosis, planning, and control of a power system. The memory is physically distributed across a network. ADEPTS instantiates a distributed process model involving synchronous message-passing. Processes interact using a rendezvous-like protocol to service requests for data. Currently, ADEPTS is implemented on a network of transputers using a Macintosh host.

ADEPTS seeks to incorporate recent innovations in object-oriented processing to the system design [1, 8]. The object-based paradigm is being used to establish a concurrent system development environment, whereby an object is viewed as an entity whose behavior is described in terms of operations it is subjected to and carries out on other objects [1]. This approach has also been used in the development of the ADEPTS

prototype in Ada [9].

Finally, the ADEPTS processing model incorporates techniques from distributed knowledge-based systems. It includes the idea of a small, loosely-coupled group of semi-autonomous managers (defined below) cooperating to continuously maintain the proper distribution of electrical power to a set of loads. The organization model loosely chosen as the paradigm for ADEPTS is the mission control model, wherein a committee of experts cooperate to maintain the proper flow of power through the spacecraft.

For clarity, the following terminology is used to describe the processing components of ADEPTS. ADEPTS is viewed as a network of managers. A manager is a processing unit which has been allocated a role (set of tasks, expertise). A role may consist of some expertise in power management (fault diagnosis, scheduling), data processing (e.g., providing an interface to the human operator) or communication and control. In order to perform a role, a manager must have direct (i.e. local, as opposed to remote) access to, and manipulate, a set of objects in memory; such objects include a database and knowledge base. A manager may also control a set of agents, which are smaller processing units dedicated to the performance of a single task among those allocated to the manager. From an architectural point of view, each manager in ADEPTS is assigned a single processor (specifically, a transputer) and some memory; it runs independently from the other managers. Agents local to the same manager run concurrently on the same processor and share memory. The result is a network of managers.

Currently, the overall power management task consists of the following roles:

* Load Management
* Fault Management
* Resource Management
* Low-level Monitor and Control Management
* User Interface

A more detailed description of each of these roles appears in [7]. The processing tasks involve a combination of data processing, knowledge management, numerical processing, and communication. The highlighted components in the above list indicate the scope of the investigation summarized in this report.

The current overall ADEPTS control architecture, including communication links, is displayed in Figure 1. Each of the managers runs on a transputer in parallel. As ADEPTS executes, math model data, representing voltage and current readings from a power system [6], is sent for evaluation and analysis to the power manager executive. Concurrently, loads are being planned and executed via the user interface and the

load manager. If the power system data indicates the presence of a fault, the fault manager is invoked to diagnose and implement corrective measures.

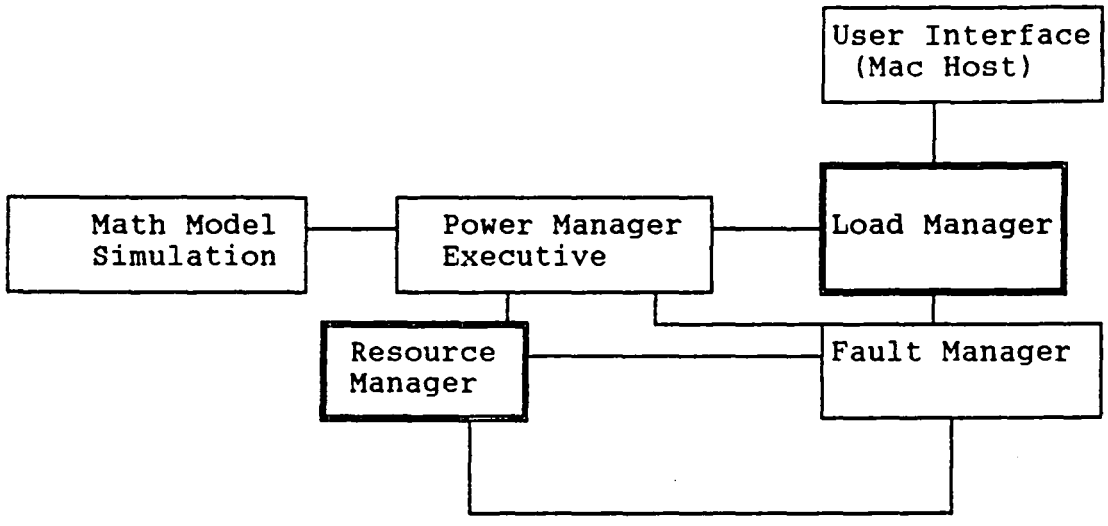


Figure 1: ADEPTS Configuration

The enhancement to ADEPTS which resulted from the project described in this report is the addition of the Load Manager and Resource Manager (represented in the figure by the double-stroke lines). The load manager was developed in order to automate the process of executing loads to a power system. The resource manager determines whether there is enough power to execute loads as planned. Finally, communication links were established between the load manager and the fault manager, the user interface, and the resource manager. Each of these components is described further in this report.

ADEPTS Software Engineering Approach

The life cycle of a real time system such as ADEPTS is complex. Figure 2 is a reproduction of (part of) a standard flow chart for an ideal real-time system life cycle [1,p.76]. ADEPTS is currently at the requirement specification phase. The result of the preliminary activity was the selection of the transputer as the architecture in which to develop ADEPTS, and the development of a small, experimental version of ADEPTS (summarized in [6]). The system requirement review was performed by initial NASA and contractor developers, and by the author of this report during last summer's research (summarized in [7]).

The software engineering philosophy of ADEPTS is that too much documentation, and not enough implementation is not a healthy approach. One reason for this is that the software and hardware technology being applied to power management has not

been extensively tested and verified. Another reason is that the developers do not have extensive programming experience on the transputer, and hence require extensive hands-on training. Thus, the developers prefer an approach which combines three activities during the current specification and review phases:

* Specification documentation
* Rapid prototyping of concepts
* Iterative development on the transputer

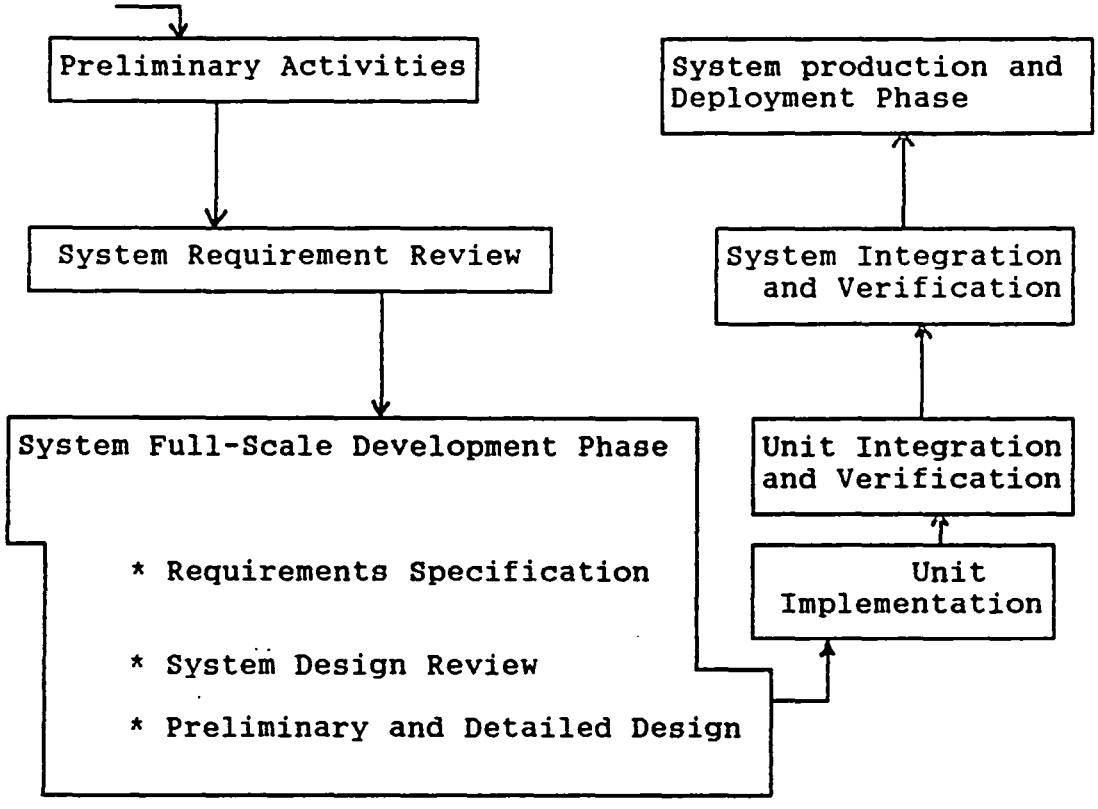


Figure 2. Real-Time System Life Cycle

The purpose of the first task is clear: to provide a detailed design and task description for ADEPTS, and to allocate subtasks to the ADEPTS developers. The purpose of the second task is to allow for a empirical verification of the overall design, and to identify any modifications or further enhancements to the design that might be required. A prototype environment for ADEPTS has been developed in Ada under Unix [9]. Finally, the purpose of the third task is to test the real time machine-dependent aspects of the design, especially the synchronization and communication mechanisms. Performing the third task also has the practical advantage of allowing the developers, many of whom may be unfamiliar with the transputer environment, with an opportunity to become more

familiar with its features.

LOAD MANAGER AND RESOURCE MANAGER DEVELOPMENT

In the following sections, we describe the steps leading to the development of the load management and resource management roles in ADEPTS. First, a set of requirement specifications for each manager is identified. Then, the data structures and algorithms used to implement these managers are described. It should be noted that the scope and complexity of each of these manager's roles prohibited an implementation of all the features specified in the requirements description in the ten weeks allocated for this activity. Thus, the discussion concludes with a list of future enhancements required for each manager.

Resource Manager Requirements Specification

The resource manager automates the power management tasks associated with system configuration, including

* Monitoring the supply of power to the system;

* Reconfiguration as a result of adding or deleting loads, or as a response to a faulty condition; and

* Load shedding as a result of too much demand for power.

The goal in developing this agent is to free the crew member and mission control personnel of routine power supply monitoring tasks, to have an effective intelligent control mechanism in the event of loss or decrease of power to the system, and to aid in the planning and execution of power system events.

To perform these tasks, a resource manager must interact with both the diagnostic and planning managers, as well as with the low-level power system monitors, who will supply it with sensor data from the main busses supplying power. The latter will allow the resource manager to determine whether a loss of power to the loads has occurred, and consequently whether shedding may be required.

The communication link with the planner (load manager) provides the resource manager with the task of filtering commands to the system for powering on loads. This filter is necessary to avoid requests for power to loads on a path which contains a fault. On the ADEPTS approach, there are two ways this scenario can be avoided. First, the load manager will be informed of faulty power system components. This will allow it to delay requests for power to such loads. The resource manager provides a second filtering mechanism; it will

prohibit situations in which a fault has occurred and identified by the fault manager, but has not been yet recorded by the load manager.

Consequently, the resource manager will act as a filter in case either there is a fault on the main bus, or there is a fault downstream of which the load manager has not been notified.

In addition to being a filter, the resource manager is also a configuration expert. It has the expertise to translate load requests into activation commands (i.e., throwing switches off or on). It also shares with the load manager information about the critical loads in the systems, those loads with redundant sources of power. If a fault on one of the paths to the critical load is detected, the resource manager will automatically reconfigure the system to the redundant path. This act may result in the lack of power to other loads, and hence the need for shedding loads, which will also be performed automatically by the resource manager. Finally, load shedding may need to be performed as a direct result of a reduction in the supply of power on the main bus; this event will be perceived by the resource manager, and the act of shedding performed.

To summarize, the resource manager has a three-fold role: as a monitor of the supply of power along the main bus, as a filter of load requests from the load manager, and as a configuration expert, which implies the ability to perform load shedding and redundancy switching for critical loads. In future enhancements to ADEPTS, each of these three sub-roles may be assigned agents.

Data Structures for ADEPTS Resource Manager

To perform its role, the resource manager maintains two lists as data: a fault list and an activation list. The fault list indicates which components of the system are faulty. The activation list indicates the status of each switch in the system (on or off). These lists are constantly being updated, as the result of its own actions (i.e. controlling the switches), as the result of diagnostic information from the fault manager, and from its own main bus sensor data indicating available power.

The expertise it has to allow it to translate load requests into switch commands resides in a configuration knowledge base. This consists of structural information about the configuration of the system, as well as load information (e.g. how much power a load needs, and whether or not it is a critical load, one with a redundant path).

ADEPTS Implementation of Resource Manager

The current ADEPTS resource manager does not perform load

shedding or redundancy reconfiguration. In pseudo-code, the current resource manager consists of the following algorithm:

```
loop
  {wait for
    fault_data; (* from fault manager *)
    load_request; (* from load manager *)
   case
    fault_data
         update_fault_list;
    load_request
       confirm_request; (* check for fault *)
       if confirmed then
        reconfigure_system; (*send switch commands*)
         update_activation_list;
  }
```

Confirm_request checks the fault list and configuration knowledge base for faulty conditions along a path to a requested load. Reconfigure_system constructs an activation command, which is a set of switch commands (on or off), which it sends to the executive, and ultimately to the math model simulation.

Future Enhancements to Resource Manager

The current ADEPTS Resource Manager implements the filtering task and part of the configuration task in fulfillment of this manager's requirements. To complete the implementation, the monitoring task must be developed, as well as the ability to perform redundancy configuration and load shedding. Finally, the load manager should receive from the resource manager an acknowledgement of its request, or, if the request cannot be met, a message indicating that the request cannot be serviced.

Load Manager Requirements Specification

The load manager is responsible for the planned control of power system events. A load manager should be flexible enough to operate either autonomously, or as an assistant to the human operator. "Offline", the load manager should manage the planning and scheduling of loads. "Online", it should contain the means to automatically cause the execution of a set of scheduled loads. In carrying out its assigned task, the load manager will free both crew member and mission control personnel from routine or time-consuming tasks related to planning and executing loads. In addition, it will provide an on-board database of load information, and procedures to assist humans in manually developing or changing load schedules.

A robust on-board load scheduler should combine both predictive and dynamic aspects. By predictive is meant the capability to plan a schedule for the future. By dynamic is meant the capability to react to unforseen events in the environment by rescheduling loads. An on-board system should also combine full autonomy with an interactive capability. It should assist the human in incrementally developing a schedule, and it should also be able to develop a full schedule from a list of loads and their constraints. The scheduling algorithm can combine traditional techniques with knowledge-based approaches. The most promising knowledge-based approach to predictive, autonomous scheduling is the constraint-based approach. This approach has been used in a number of applications, including payload scheduling for spacecraft [10]. Recently, a distributed constraint-based approach has been proposed [11].

Data Structures For Load Management

The ADEPTS load manager consists of a load table containing information about loads to be powered during a relatively short duration (one day) of time. Its permanent knowledge consists of information useful for either automatic or interactive planning and scheduling activity. Load priority, power and resource requirements, timing constraints, as well as whether the load is currently faulted, are examples of useful load information.

The load table could consist of a short term (one day) schedule of loads, or a long-term schedule, or both. The load event should consist of a load identifier, and the interval(s) during which it is to be powered, comprising a start time and end time. (The load table can be displayed to the user as a table of time intervals). Thus, for example, Payload L3 has been scheduled once during this time frame, during the interval starting at mission time 2 and extending to mission time 9.

The load table is implemented in ADEPTS as a linked list of load information. This information is of two kinds: status information and occurrence information. Status information is a record implemented in C as follows:

```
typedef struc
 { short activation status;
   short fault status;
   tslot timeslot; }
 loadrecord;
```

Activation status (OFF or ON) indicates whether a load currently has power. Fault status (FAULTED, HEALTHY) indicates whether there is a fault somewhere in the line leading to the load.

Occurrence information contains the start and end times of each scheduled occurrence of the load. This allows a load to have multiple, periodic occurrences during the same time frame. The C record for an occurrence has the following declaration:

```
typedef struc
 { ulong startTime ;
   ulong endtime ;
   tslot timeslot;(* pointer to next load
                     occurrence*)
 }
   tslotrecord;
```

These data structures should be enhanced in future work to include other information useful in the scheduling or rescheduling of loads. For example, load priority information, resource requirements (power, instruments, crew, etc.) could be declared as additional slots. Since the current load manager does not engage in any scheduling activities, it was not deemed necessary to include these fields in the definitions.

Load Manager Functional Specification

The load manager role can be classified into three subtasks:

* Monitoring/executing events from the load table
* Updating the load table
* Scheduling and rescheduling loads

These tasks can be viewed as independent; hence, as involving separate agents. On the other hand, each operates on the same data structure, the load map. Hence, for future ADEPTS enhancements, it should be possible to perform these tasks concurrently, but probably using one transputer.

The communication links required by load manager are to the User Interface and Resource Manager (Figure 1). The user interface link allows the user (crew member or mission control personnel) to update the load table as required. It also allows for the communication and display of load table and other load information to the human operator. The resource manager link allows for the coordination of load execution activity. As noted, the resource manager will not cause a load to be executed if a fault has been identified by the fault manager along a path leading to this load, unless a redundant path exists to this load, and there is enough redundant power to allow the load to be executed. To avoid the useless communication of load execution information between the load

manager and resource manager, the load manager can receive current fault information from the resource manager. This allows the load manager to update its load table so that loads along the path of a fault will not be executed until the fault has been corrected.

C Implementation of Load Manager Functions

The main procedure of the load manager currently consists of calls to initialize a time map and the local clock, and a continuous loop for monitoring and executing loads:

```
void main
{
      InitializeTimeBase ();
      InitializeLoadRecord ();

     do
      CheckSchedule ();
     while (DoCommand ());
}
```

InitializeTimeBase requests the real time from the main clock, and converts it into a "mission time" framework, where the 0th time point is the onset of the mission. This time measure corresponds to the metric used in the load map; i.e., each load map starts at 0 mission time. This simplified time framework suffices for the purposes of this "not fully real-time" version of ADEPTS. Future versions will demand more complex procedures for representing real time and mission time.

The do loop in the main procedure handles both the communication and time map monitoring. DoCommand is a procedure for input channel communication. The input will be in the form of schedule information, either from a file, or through user commands from the keyboard. CheckSchedule is the monitor and load execution procedure. It performs a traversal of the time map to search for loads that need to change their status (from on to off or visa versa). To match global time with mission time, a function for converting one to the other is employed. On the Mac, global time is depicted as number of seconds that have elapsed since midnight on January 1, 1904. This value is translated into mission time within Checkschedule.

If a load status needs to be changed, load information in the form of a load identifier and a command, is communicated to the resource manager.

Possible Enhancements to Load Manager

A number of enhancements is required for Load Manager to

meet design specifications. In sum, these include:

* Development of Predictive Scheduler
* Development of Enhanced Dynamic Scheduler
* A Load Map Display Mechanism
* A Finer Partitioning of Management Tasks
* A More Object-oriented Approach to Knowledge Representation
* Development of a More Interactive Approach to Scheduling (More User Friendly Control)

FUTURE OVERALL ADEPTS ENHANCEMENTS

In addition to the enhancements to the load and resource managers recommended above, additional development efforts have been identified as the result of this summer's activity. In this section, these are summarized.

Certain limitations of the Macintosh/transputer network have been identified and should be addressed in the future. First, major limitations of the transputer are due to its message passing protocol. In ADEPTS, the transputer hardware and software forces a synchronous exchange between processes. This means that the execution of an output command by a process i is delayed until the receiving process j is ready to execute a read operation, and visa versa. There are potentially serious problems as a result of imposing synchronism on processes. One of the most serious is the absence of fault-tolerance of transputer networks; if a receiving process fails, each of the processes which communicate with it will eventually fail as well. It should be noted that a form of asynchronism can be achieved through employing a timed-out mechanism in the communication. As ADEPTS develops, this mechanism should be considered more fully.

Additional limitations arise as a result of the interaction between the Macintosh host and the transputers. The host computer is required to simultaneously process user input and transputer data. The programmer is required to establish the protocol for servicing these requests. If the transputer is given full priority, the user may be indefinitely forced to wait on his or her request. On the other hand, if the user is given priority, the transputer is forced to wait until its request can be serviced, which means that its other processing tasks may be thwarted. The current solution employed in ADEPTS allows for fairness in the servicing of requests from the user and the transputer. As a result, however, the performance of the transputers may still be affected, especially if it is assumed that the user will use the system often.

Part of the problem is that the current Machintosh operating system (6.0) does not support multitasking

capabilities. As a result, the programmer is forced to ensure fairness and liveness properties of the system. Future versions of the Macintosh operating system (7.0) will support multi-tasking, and hence improve the real-time capabilities of the system, and free the programmer from the responsibilities of enforcing fairness and liveness properties. It behooves the developers of ADEPTS to address the problem of the host/transputer interaction in more detail.

The final set of enhancements to be recommended for future ADEPTS development activity address the furthering of the effort of establishing an overall design for a real time ADEPTS system.

In general, ADEPTS can be viewed as having four layers:

* The Interface layer
* The Communication and Control layer
* The Data and Numerical Processing layer
* The Expert System layer

Each of these layers requires a more detailed set of requirements than is currently available. The interface layer requires specifications for the user interface, as well as the lower-level interfaces to the system hardware. The communication and control layer demands a set of specifications for communication among managers, and among agents within a manager, as well as deciding on a more effective global control mechanism for the performance of management tasks. The data and numerical processing layer defines the data to be used and communicated among managers, as well as the algorithms for processing the data. Finally, the expert system layer addresses problems of knowledge base representation and management, handling uncertainty, and heuristic reasoning.

SUMMARY

The ADEPTS project seeks to integrate expert system technology with parallel and distributed computing for the purpose of developing an autonomous system for spacecraft power management. The primary focus of this summer's activity was on the means of increasing the autonomy of ADEPTS by developing techniques for the automatic scheduled execution of payloads, as well as a means of coordinating planning with limitations in power resources. Implementing these techniques also aided the developers in evaluating the transputer as a parallel computing hardware for ADEPTS. A secondary focus of this summer's work was to continue the effort of developing a set of design specifications for a more real time version of ADEPTS. The overall result of this effort has been a more precise and detailed description of the ADEPTS control mechanism and agent roles.

REFERENCES

1. Levi, S. and Agrawala, A. Real Time System Design. McGraw-Hill, 1990.

2. Hughes, J. Object-Oriented Databases. Prentice Hall, 1991.

3. Van Leeuwen, J. Handbook of Theoretical Computer Science, Vol. B: Formal Models and Semantics. The MIT Press/Elsevier, 1991.

4. Emerson, E.A., Temporal and Modal Logic. In [3], Chapter 16.

5. Lamport, L., and Lynch, N. Distributed Computing: Models and Methods. In [3], Chapter 18.

6. Sandman, T., Advanced Electrical Power Techniques for Space Systems, Phase I Report. Document JSC-24745, October 1990.

7. Morris, R., A Design for an Intelligent Monitor and Controller For Space Station Electrical Power Using Parallel and Distributed Problem Solving. Final Report, NASA/ASEE Summer Faculty Fellowship, Johnson Space Center, 1990.

8. Chin, R., and Chanson, S., Distributed Object-Based Programming Systems. ACM Computing Surveys, Vol. 23, No. 1, March 1991, pp. 91-124.

9. Morris, R., and Baggs, R. Distributed Intelligence Monitoring and Control. NASA-JSC Final Contractor Report, T-6956R, 1991.

10. Touchton, R.A. Common Module Dynamic Payload Scheduler Expert System. 21st Intersociety Energy Conversion Conference, Washington, DC, American Chemical Society, 1986.

11. Rzevski, G. Distributed Intelligent Knowledge Based Systems for Cell Scheduling and Control. ISATA 19th International Symposium on Automotive Technology and Automation, Monaco, 1988.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1991	Contractor Report

4. TITLE AND SUBTITLE
National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program - 1991 Volume 1

5. FUNDING NUMBERS
NGT 44-001-800

6. AUTHOR(S)
William A. Hyman and Stanley H. Goldstein, Editors

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Texas A&M University
College Station, Texas

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
University Programs Office
Lyndon B. Johnson Space Center
Houston, Texas

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA CR 185670

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified/Unlimited

Star Category 99

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The 1991 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by Texas A&M University and JSC. The program at JSC, as well as the programs at other NASA Centers, was funded by the Office of University Affairs, NASA Headquarters Washington, D.C. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are (1) to further the professional knowledge of qualified engineering and science faculty members; (2) to stimulate an exchange of ideas between participants and NASA; (3) to enrich and refresh the research and teaching activities of participants' institutions; and (4) to contribute to the research objective of the NASA Centers.

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project in collaboration with a NASA JSC colleague. This document is a compilation of the final reports on the research projects performed by the faculty fellows during the summmer of 1991. Volume 1 contains reports 1 through 14, and Volume 2 contains reports 15 through 27.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

End of Document